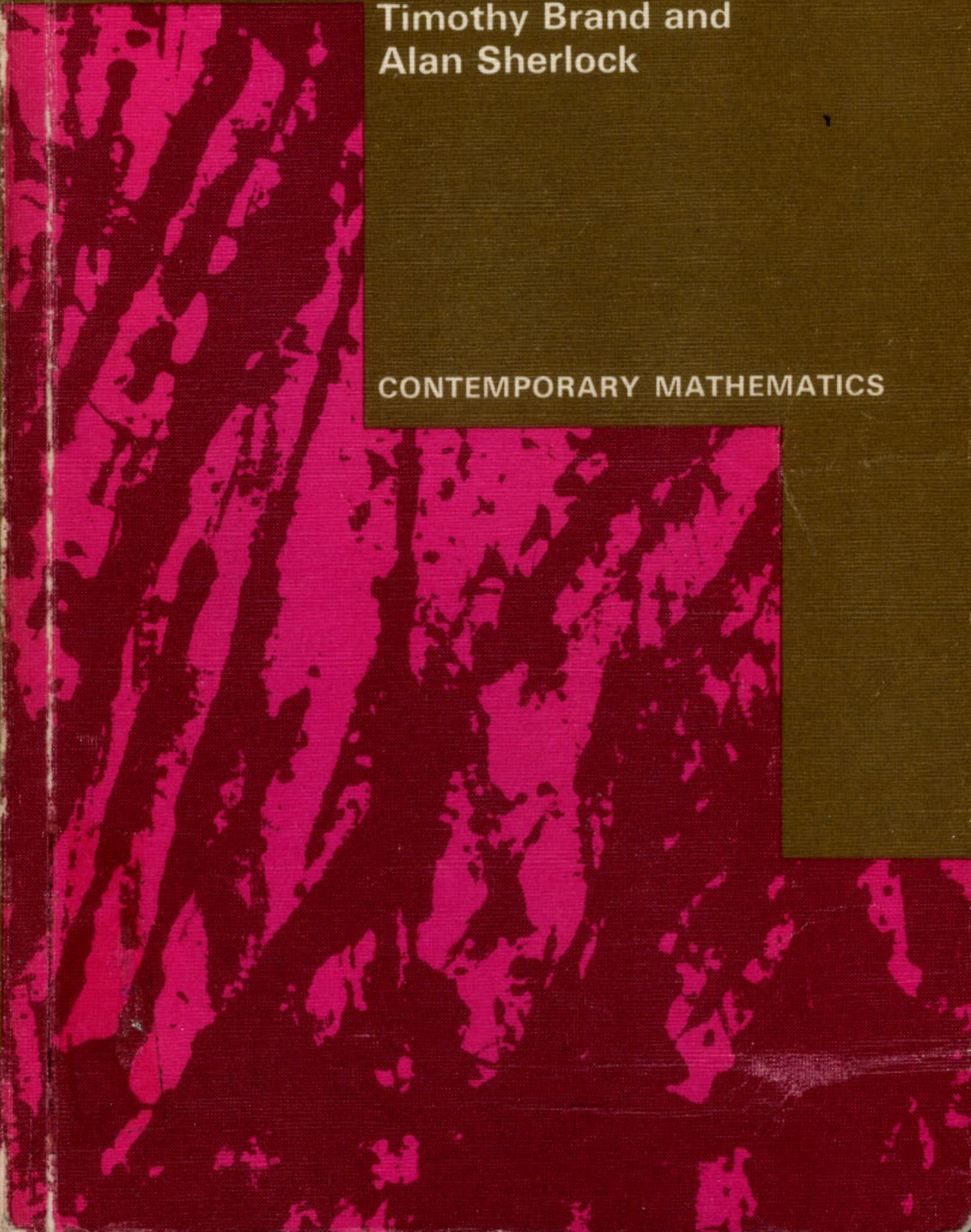
MATRICES:
Pure and Applied

Timothy Brand and
Alan Sherlock

CONTEMPORARY MATHEMATICS



Brand &
Sherlock

MATRICES: Pure and Applied

Arnold

Matrices: Pure and Applied

The Contemporary Mathematics Series comprises:
Matrices: Pure and Applied Timothy Brand and Alan Sherlock
An Elementary Introduction to Operational Research
J. Q. Thompson
Analogue Computers Michael Brand and Timothy Brand

Further texts on Vectors, Structure and
other subjects are in development

Dedicated to
Geoffrey who began it all

Contemporary Mathematics

General Editor: Alan J. Sherlock

MATRICES : Pure and Applied

Timothy Brand and Alan Sherlock

EDWARD ARNOLD
London

© Timothy E. Brand and Alan J. Sherlock 1970
First published 1970

SBN: 7131 2246 3

PRESTON
POLYTECHNIC
53040
C
512.943 BRA

4/s
5th copy

Printed in Great Britain by
Western Printing Services Ltd
Bristol

Editor's Preface

The revolution in the teaching of mathematics at schools and colleges is now well established. Not only have the *methods* of teaching changed considerably, but so have the *topics* which are taught. This is as it should be—the needs of our future mathematicians and scientists are constantly changing in the light of the technical developments of the twentieth century. It seems likely that the frontiers of science will be pushed further and further back in the coming years and the mathematics required must certainly keep abreast of these changes. If these developments are to continue rapidly, it is necessary for the students to meet and understand new ideas whilst in their last years at school and their first at college and university.

The 'Contemporary Mathematics' series of books and booklets has been written after three years' thought and experiment, and has been designed for these major purposes:

1. To produce texts to cover the new requirements of the major school mathematics projects ('A' and 'S' level in Britain). Texts will provide the mathematical developments and cover the important applications (e.g. *Matrices: Pure and Applied*).
2. To provide an introduction to topics which are both contemporary and important, showing also how these might best be taught (e.g. *An Elementary Introduction to Operational Research*).

Many schools are experimenting with first examination ('O' level) work, but in senior classes ('A' level) the experiments are just developing.

It would seem sensible to develop and continue a number of the first examination ('O' level) topics—for example, Boolean Algebra, Matrices and Computers all combine in Boolean Matrices (the key to many computer circuits). Matrices, Vectors, Numerical Methods, Analogue and Digital Computing and Probability and Statistics are part of every new course and are to be included.

It is hoped that these books will help students to appreciate at least a few of the current, most pressing applications which must soon be part of every mathematician's repertoire.

London
1969

A. J. S.

Authors' Preface

This book is intended as an extension and continuation of *Matrices 1* and *Matrices 2* (both by G. Matthews) at Sixth Form level. The book contains a short revision of 'O' level matrix work but is, in fact, completely self-contained. The first section (Chapters 1–5) is an investigation of some aspects of matrix algebra and covers all the known 'A' level syllabuses. The purpose of this section is to provide the mathematical tools which will be of most use later on when dealing with the applications. The second section (Chapters 6–9) is an attempt to give an insight into some of the many and varied applications of matrices.

It is impossible to include a complete list of applications as this would be far too large, and undoubtedly some interesting topics have been excluded. However, the selection given has been taken from as wide a field as possible and we hope will be of great interest.

Acknowledgements

We should like to express our grateful thanks to Geoffrey Matthews and Trevor Fletcher for their initial inspiration and for some of the ideas; also to the many authors and lecturers from whom we have gleaned interesting topics and other ideas, especially to Alastair McDougall in Zambia.

We should also like to express our appreciation of the help given to us by our colleagues at St Dunstan's—Bill Wade, Quentin Thompson and Tim Pratt, all of whom have helped by their willing co-operation in teaching the material and by their forthright comments. Thanks also to our pupils for many suggestions, and especially to Andrew Ellis for his article on complex matrices and for his help in inventing questions.

Finally, we should like to thank the Oxford and Cambridge Joint Examination Board, the University of Cambridge, the University of Leicester, the M.E.I. Schools and the Associated Examination Board for permission to use some of their questions.

London
1969

T. E. B.
A. J. S.

Bibliography

This book is built on the foundations provided by the St Dunstan's College Booklets (Contemporary School Mathematics published by Edward Arnold Ltd). In particular:

BAILEY, C. A. R. *Sets and Logic 1* and *Sets and Logic 2.*

BRAND, T. E. and WADE, D. W. M. *Exercises 2.*

MATTHEWS, G. *Matrices 1* and *Matrices 2.*

Many other books and articles were found most useful. The following list is not complete and some further references have been given in the text.

BIRKHOFF, G. and MCLANE, S. *A Brief Survey of Modern Algebra.* Third edition. Macmillan, London, 1965.

BOWMAN, F. *An Introduction to Determinants and Matrices.* English Universities Press, London, 1962.

CAMPBELL, H. G. *Matrices with Applications.* Appleton-Century-Crofts, New York, 1968.

EDGE, G. M. and BARRON, I. M. *Understanding Boolean Algebra.* Cambridge Consultants Ltd. (reprinted from *Instrument and Control Engineering*).

FINKBEINER, D. T. *Introduction to Matrices and Linear Transformations.* Freeman, London, 1966.

FLEGG, H. G. *Boolean Algebra and Its Applications.* Blackie, Glasgow, 1964.

FLETCHER, T. (Editor).* *Some Lessons in Mathematics.* Cambridge University Press, 1964.

HEADING, J. *Matrix Theory for Physicists.* Longmans Green, London, 1958.

FRAZER, R. A., DUNCAN, W. J. and COLLAR, A. R. *Elementary Matrices.* Cambridge University Press, 1939.

HLAWICZKA, P. *Matrix Algebra for Electronic Engineers.* Iliffe, London, 1965.

KEMENY, J. G. and SNELL, J. L. *Finite Markov Chains.* Van Nostrand, New York and London, 1960.

---

* Mr Fletcher's unpublished 'Teachers' Notes on Matrices' have also been of great help.

KEMENY, J. G., SNELL, J. L. and THOMPSON, G. L. *Introduction to Finite Mathematics*. Prentice-Hall, New Jersey and London, 1966.

MAXWELL, E. A. *Algebraic Structure and Matrices*. Cambridge University Press, 1965.

NEILL, H. and MOAKES, A. J. *Vectors, Matrices and Linear Equations*. Oliver & Boyd, Edinburgh, 1967.

PIPES, L. *Matrix Methods for Engineers*. Prentice-Hall, New Jersey and London, 1963.

PRENTIS, J. M. and LECKIE, F. A. *Mechanical Vibrations*. Longmans Green, London, 1962.

ROGERS, F. E. *Topology and Matrices in the Solution of Networks*. Iliffe, London, 1965–7.

SEARLE, S. R. *Matrix Algebra for Biological Sciences*. Wiley, New York and Chichester, 1966.

STEPHENSON, G. *An Introduction to Matrices, Sets and Groups*. Longmans Green, London, 1965.

WHITESITT, J. E. *Boolean Algebra and Its Applications*. Addison Wesley, London, 1961.

Contents

Editor's Preface v
Authors' Preface vii
Bibliography ix
1 Matrices and Determinants: An Introduction 1
2 Elementary Matrices: Partitioned Matrices 20
3 Linear Equations 40
4 Matrix Transformations 82
5 Powers of Matrices 117
6 Karnaugh Maps and Associated Matrices 145
7 Boolean Matrices 161
8 Matrices and Networks 174
9 Further Situations in which Matrices are Used
  I An example from the theory of numbers 187
  II Recurrence relations 194
  III Difference equations 198
  IV Solution of differential equations 201
  V Four-terminal networks 217
  VI Affine transformations 223
  VII Matrices and the theory of games 228
  VIII Knots 234
  IX Matrices with complex entries 238
Answers 243

1

Matrices and Determinants: An Introduction

1.1 A *matrix* is a rectangular array of m rows and n columns and is [illegible] by

$$M = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdot & \cdot & \cdot & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdot & \cdot & \cdot & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdot & \cdot & \cdot & a_{3n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & a_{m3} & \cdot & \cdot & \cdot & a_{mn} \end{pmatrix}$$

The matrix M above is said to be of order $m \times n$ because it has m rows and n columns. Each element a_{ij} or entry of the matrix belongs to a given set,* where i refers to the row and j to the column, e.g. a_{23} would be the element in row 2, column 3.

SOME MATRICES

A matrix that is of order $n \times n$ is a *square matrix*, i.e. has the same number of rows and columns.
A matrix that is of order $1 \times n$ is a *row vector*, e.g. $(1 \quad 2 \quad 3 \ldots n)$.
A matrix that is of order $m \times 1$ is a *column vector*. For example,

$$\begin{pmatrix} 1 \\ 2 \\ 3 \\ \cdot \\ \cdot \\ \cdot \\ m \end{pmatrix}$$

---

* Initially the set of real numbers R, but the results are true for the set of complex numbers. Later, other sets will be considered, e.g. matrices with Boolean entries.

A matrix with all elements zero is called a *zero matrix*. For example,

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad (0 \quad 0), \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

EQUALITY OF MATRICES

Two matrices are said to be equal if (*a*) they are of the same order, and (*b*) their corresponding elements are equal.

ADDITION OF TWO MATRICES

Two matrices A and B can be added if, and only if, they are of the same order. The sum $A + B$ is obtained by adding corresponding elements. For example,

if $$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{pmatrix}$$

then $$A + B = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \\ a_{31} + b_{31} & a_{32} + b_{32} \end{pmatrix}$$

MULTIPLICATION OF A MATRIX BY A SCALAR

If $k \in R$ and M is the matrix defined earlier then kM is the matrix formed by multiplying every element a_{ij} of M by k. For example,

$$2\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} = \begin{pmatrix} 2a & 2b & 2c \\ 2d & 2e & 2f \end{pmatrix}$$

Exercise 1.1

1. What is the order of the following matrices?

(*a*) $\begin{pmatrix} 9 & 2 \\ 7 & 4 \end{pmatrix}$ (*b*) $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ (*c*) (1)

(*d*) $(p \quad q)$ (*e*) $\begin{pmatrix} 2b & 3c \\ d & 2b \\ 0 & 5a \end{pmatrix}$ (*f*) $\begin{pmatrix} 2 & 1 & 0 \\ 3 & 2 & 5 \end{pmatrix}$

(*g*) $(4 \quad 2 \quad 3 \quad 1)$ (*h*) $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ (*i*) $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$

2. If A, B, C are all matrices of order $m \times n$, and the elements of each of them belong to R, show that

(i) $A + B = B + A$ (the commutative law)

(ii) $(A + B) + C = A + (B + C)$ (the associative law)

3. A is an $m \times n$ matrix and B is the matrix $(-1)A$, find the matrix $A + B$. Use these results to formulate a definition of matrix subtraction.

4. If A is an $m \times n$ matrix and O is the $m \times n$ zero matrix, what is $A + O$?

5. Prove that the set of $m \times n$ matrices with elements belonging to R form a commutative group under addition.

Multiplication of two matrices

1.2 If A is a matrix of order $m \times n$, and B a matrix of order $p \times q$, the product AB is only defined if $n = p$. In this case AB is a matrix of order $m \times q$. Matrix multiplication is illustrated by the example.*

$$A = \begin{pmatrix} \boxed{a_{11} \quad a_{12}} \\ a_{21} \quad a_{22} \\ a_{31} \quad a_{32} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} \boxed{\begin{matrix} b_{11} \\ b_{21} \end{matrix}} & \begin{matrix} b_{12} \\ b_{22} \end{matrix} \end{pmatrix}$$

$$\Rightarrow AB = \begin{pmatrix} \boxed{a_{11}b_{11} + a_{12}b_{21}} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{31}b_{12} + a_{22}b_{22} \\ a_{31}b_{11} + a_{32}b_{21} & a_{31}b_{12} + a_{32}b_{22} \end{pmatrix}$$

The ringed elements show how one element is formed, for

$$(a_{11} \quad a_{12})\begin{pmatrix} b_{11} \\ b_{21} \end{pmatrix} = (a_{11}b_{11} + a_{12}b_{21})$$

In this example BA does not exist.

Exercise 1.2

1. If
$$A = (1 \quad 2 \quad 3),$$
$$B = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} \quad \text{and} \quad C = \begin{pmatrix} x & y & z \\ p & q & r \end{pmatrix}$$

state which of the following products exist. Work out those which do.

(*a*) AB (*b*) BA (*c*) AC

(*d*) CA (*e*) BC (*f*) CB

---

* This method of matrix multiplication can be justified by repeated application of linear transformations (see p. 91).

2. Is matrix multiplication commutative? If not give a counter example?

3. If A is of order $m \times n$ and B is of order $p \times q$, what is the condition for both AB and BA to exist?

4. If A and B are defined as in **1.2** and

$$C = \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \end{pmatrix}$$

show that $(AB)C = A(BC)$. Can you extend this result to show that all matrix multiplication is associative?

5. By considering three matrices P, Q and R, not all square but of suitable orders, investigate whether $P(Q + R) = PQ + PR$.

6. Find the conditions that must be imposed on the order of two matrices for them to be commutative under multiplication, i.e. $AB = BA$.

7. Find two matrices A and B such that $AB = O$ but such that neither $A = O$ nor $B = O$.

The transpose matrix

1.3 If

$$M = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

then the transpose is that matrix whose rows are the columns of M. It is written as

$$M' = \begin{pmatrix} a & d & g \\ b & e & h \\ c & f & i \end{pmatrix}$$

Another example. If

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \text{ then } A' = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}$$

Exercise 1.3

1. If

$$A = \begin{pmatrix} a & b & c \\ d & e & g \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} p & q \\ r & s \\ t & u \end{pmatrix}$$

show that

$$(AB)' = B'A'$$

2. If

$$A = \begin{pmatrix} a & h & g \\ h & b & f \\ g & f & c \end{pmatrix} \quad \text{and} \quad \mathbf{x} = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

find $\mathbf{x}'A\mathbf{x}$.

3. A square matrix is said to be *stochastic* if all the elements are non negative and the sum of the elements in each row is 1. Show that the product of two 2 × 2 stochastic matrices is stochastic.

There is a row vector $R = (x \ y)$ such that $x + y = 1$, associated with a stochastic matrix

$$S = \begin{pmatrix} 1 - a & a \\ b & 1 - b \end{pmatrix}$$

and satisfying the relation $RS = R$. Find x and y in terms of a and b. Hence find the row vector R associated with the matrix

$$\begin{pmatrix} \frac{1}{4} & \frac{3}{4} \\ \frac{5}{6} & \frac{1}{6} \end{pmatrix}$$

(*SMP Additional*)

DETERMINANTS

1.4 The *determinant* $|A| = \begin{vmatrix} a & b \\ c & d \end{vmatrix}$ of the 2 × 2 matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is an alternative notation for the expression $ad - bc$.

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}$$, the determinant of a 3 × 3 matrix,

is defined as

$$a\begin{vmatrix} e & f \\ h & i \end{vmatrix} - b\begin{vmatrix} d & f \\ g & i \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix}$$

1.5 Note that in this expansion by the elements of the first row, the 2 × 2 determinants are obtained by 'covering up' the row and column in which each first row element appears, i.e.

$$a\begin{vmatrix} \cancel{a} & \cancel{b} & \cancel{c} \\ \cancel{d} & e & f \\ \cancel{g} & h & i \end{vmatrix} - b\begin{vmatrix} \cancel{a} & \cancel{b} & \cancel{c} \\ d & \cancel{e} & f \\ g & \cancel{h} & i \end{vmatrix} + c\begin{vmatrix} \cancel{a} & \cancel{b} & \cancel{c} \\ d & e & \cancel{f} \\ g & h & \cancel{i} \end{vmatrix}$$

In the same way, the determinant

$$\begin{vmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & o & p \end{vmatrix}$$

is defined as

$$a\begin{vmatrix} f & g & h \\ j & k & l \\ n & o & p \end{vmatrix} - b\begin{vmatrix} e & g & h \\ i & k & l \\ m & o & p \end{vmatrix} + c\begin{vmatrix} e & f & h \\ i & j & l \\ m & n & p \end{vmatrix} - d\begin{vmatrix} e & f & g \\ i & j & k \\ m & n & o \end{vmatrix}$$

Determinants of higher order square matrices can be defined in a similar manner.

Exercise 1.4

1. Evaluate the following determinants:

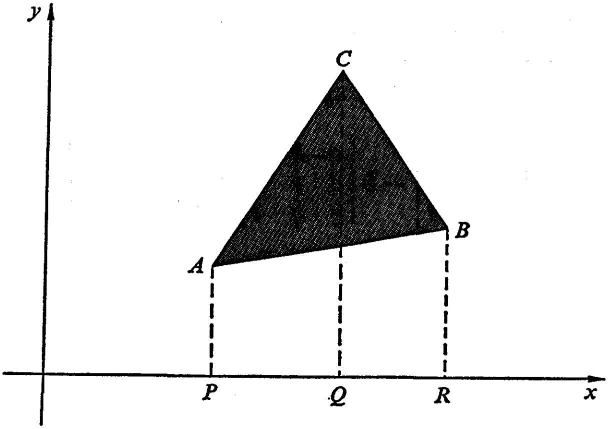
(*a*) $\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}$ (*b*) $\begin{vmatrix} 1 & -1 \\ 1 & -1 \end{vmatrix}$

(*c*) $\begin{vmatrix} 8 & -5 \\ 44 & 2 \end{vmatrix}$ (*d*) $\begin{vmatrix} a & b \\ b & a \end{vmatrix}$

(*e*) $\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 4 \\ 3 & 2 & 1 \end{vmatrix}$ (*f*) $\begin{vmatrix} -1 & 0 & 3 \\ 2 & 1 & 4 \\ -2 & -3 & -1 \end{vmatrix}$

(*g*) $\begin{vmatrix} 1 & 2 & 3 \\ -1 & 0 & 4 \\ 1 & 2 & 3 \end{vmatrix}$ (*h*) $\begin{vmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{vmatrix}$

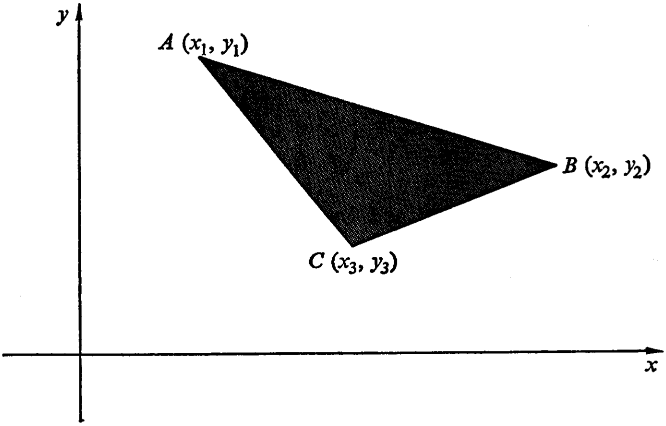
2. (*a*)



By considering trapezia such as *PACQ*, or otherwise, show that the area of the triangle *ABC* with vertices at $A(x_1, y_1)$, $B(x_2, y_2)$, and $C(x_3, y_3)$ is given by

$$\frac{1}{2}\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}$$

(*b*) Find an expression for the area of the triangle as given in the figure below



(*c*) Show that

$$\begin{vmatrix} x & y & 1 \\ a & b & 1 \\ c & d & 1 \end{vmatrix} = 0$$

is the equation of a straight line passing through two distinct points (a, b) and $(c, d.)$

3. If M is the matrix

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

show that $|M| = |M'|$.

4. If M is the matrix defined in question 3, and M_1 is the matrix obtained by interchanging any two rows of M, show that

$$|M| = -|M_1|$$

5. Use the results of the last two questions to find (without expansion) the value of the following determinants, given that

$$\begin{vmatrix} 1 & -1 & 2 \\ 3 & 4 & -5 \\ 0 & -2 & -3 \end{vmatrix} = -43$$

(*a*) $\begin{vmatrix} 0 & -2 & -3 \\ 3 & 4 & -5 \\ 1 & -1 & 2 \end{vmatrix}$ (*b*) $\begin{vmatrix} 2 & 1 & -1 \\ -5 & 3 & 4 \\ -3 & 0 & -2 \end{vmatrix}$

6. (*a*) If

$$M = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

and M_1 is the matrix obtained by multiplying any row of M by λ, show that $\lambda|M| = |M_1|$.

(*b*) Use the result of question 3 to show that if M_2 is found by multiplying any column of M by λ then $\lambda|M| = |M_2|$.

(*c*) If

$$\begin{vmatrix} 1 & -1 & 2 \\ 3 & 4 & -5 \\ 0 & -2 & -3 \end{vmatrix} = -43$$

find, without expansion, the value of the following determinants:

(i) $\begin{vmatrix} 10 & -10 & 20 \\ 3 & 4 & -5 \\ 0 & -2 & -3 \end{vmatrix}$ (ii) $\begin{vmatrix} 6 & -2 & 4 \\ 9 & 4 & -5 \\ 0 & -2 & -3 \end{vmatrix}$

(iii) $\begin{vmatrix} 1 & 1 & -2 \\ -3 & 4 & -5 \\ 0 & -2 & -3 \end{vmatrix}$

(*d*) Show that

$$\begin{vmatrix} 60 & 24 & 84 \\ 15 & 1 & 14 \\ 30 & 6 & 21 \end{vmatrix} = 2 \times 2 \times 3 \times 3 \times 5 \times 7 \times \begin{vmatrix} 1 & 2 & 1 \\ 3 & 1 & 2 \\ 2 & 2 & 1 \end{vmatrix}$$

$$= 2^2 \times 3^3 \times 5 \times 7$$

7. (*a*) Show that

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = \begin{vmatrix} a+\lambda d & b+\lambda e & c+\lambda f \\ d & e & f \\ g & h & i \end{vmatrix} \quad \text{where } \lambda \in R$$

(*b*) Use the results of question 4 to show that a constant multiple of any one row can be added to any other row without altering the value of the determinant. Use the result of question 3 to show that this is true also for the columns.

(*c*) Evaluate

(i) $\begin{vmatrix} 51 & 61 & 71 \\ 5 & 6 & 7 \\ 1 & 1 & 1 \end{vmatrix}$ (ii) $\begin{vmatrix} 16 & 19 & 23 \\ 15 & 18 & 22 \\ 13 & 17 & 20 \end{vmatrix}$

Some properties of a 3×3 determinant

1.6 The following properties were illustrated in the previous exercise.

1. The value of the determinant is unaltered by interchanging its rows and columns.

2. Interchanging any two adjacent rows (or columns) changes the sign of the determinant.

3. If all the elements of any row (or column) are multiplied by a constant λ then the value of the determinant is multiplied by λ.

4. If to the elements of any one row (or column) a multiple of any other row (or column) is added, the value of the determinant remains unaltered.

The minors of a 3×3 determinant

1.7 Consider the determinant

$$\Delta = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}$$

Covering up the row and column in which the element *a* appears, the determinant $\begin{vmatrix} e & f \\ h & i \end{vmatrix}$ is obtained.

This determinant is known as the minor of *a*. Now

$$\Delta = a\begin{vmatrix} e & f \\ h & i \end{vmatrix} - b\begin{vmatrix} d & f \\ g & i \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix}$$

$$= a \times (\text{the minor of } a) - b \times (\text{the minor of } b) + c \times (\text{the minor of } c)$$

1.8 To expand by another row (or column), consider for example

$$\begin{aligned}\Delta &= a(ei - fh) - b(di - fg) + c(dh - eg)\\ &= aei - afh - bdi + bfg + cdh - ceg\\ &= -bdi + cdh + aei - ceg - afh + bfg\\ &= -d(bi - ch) + e(ai - cg) - f(ah - bg)\\ &= -d\begin{vmatrix} b & c\\ h & i\end{vmatrix} + e\begin{vmatrix} a & c\\ g & i\end{vmatrix} - f\begin{vmatrix} a & b\\ g & h\end{vmatrix}\\ &= -d \times (\text{minor of } d) + e \times (\text{minor of } e) -\\ &\qquad\qquad f \times (\text{minor of } f)\end{aligned}$$

This is the expansion of the determinant by the second row.

To expand by any other row or column is now easy, once the rule for the signs has been obtained. By regrouping the terms as before it is easily seen that the rule for signs is

$$\begin{vmatrix} + & - & +\\ - & + & -\\ + & - & +\end{vmatrix}$$

For example, the expansion of Δ by the 3rd column is now

$$c\begin{vmatrix} d & e\\ g & h\end{vmatrix} - f\begin{vmatrix} a & b\\ g & h\end{vmatrix} + i\begin{vmatrix} a & b\\ d & e\end{vmatrix}$$

The cofactors of a 3×3 determinant

1.9 If Δ is given as above,

The cofactor of element a is $A = +\begin{vmatrix} e & f\\ h & i\end{vmatrix} = +$ (the minor of a)

and

The cofactor of element b is $B = -\begin{vmatrix} d & f\\ g & i\end{vmatrix} = -$ (the minor of b)

The cofactors of the other elements can be found similarly using the rule of signs above.

The expansion of a determinant can now be obtained for any row or column in terms of its cofactors, e.g.

$$\Delta = aA + bB + cC = bB + eE + hH$$

Factorization of determinants

1.10 Symmetrical determinants can easily be factorized by the use of the Remainder Theorem (if two rows or two columns of a determinant are equal then the value of the determinant is zero) and an appeal to symmetry.

Example

Factorise

$$\begin{vmatrix} x & y & z \\ x^2 & y^2 & z^2 \\ yz & zx & xy \end{vmatrix}$$

(i) If $x = y$ the first two columns are identical. Hence $x - y$ is a factor. By symmetry $y - z$ and $z - x$ are also factors.

(ii) The leading diagonal is

$$\begin{matrix} x & & \\ & y^2 & \\ & & xy \end{matrix}$$

which would give on multiplication x^2y^3, i.e. something of order 5.

Now we already have factors of order 3 $\{(x - y)(y - z)(z - x)\}$ and we require a symmetrical factor of order 2. There are only two possible—either $x^2 + y^2 + z^2$ or $xy + yz + zx$ (or a linear combination of these). By comparing the possibilities we see that the factors must be

$$K(x - y)(y - z)(z - x)(xy + yz + zx) \text{ where K is a constant}$$

and K is seen to be one.

It is not always possible to consider only the leading diagonal as on occasion this term may cancel.

The simple symmetrical functions are for example

$$x + y + z; \quad x^2 + y^2 + z^2; \quad xy + yz + zx.$$

Exercise 1.5

1. Evaluate

(i) $\begin{vmatrix} 1 & 15 & -97 \\ 0 & 84 & 79 \\ 0 & 2 & 1 \end{vmatrix}$ (ii) $\begin{vmatrix} 27 & 85 & 94 \\ 1 & -1 & 0 \\ 4 & 2 & 100 \end{vmatrix}$

(iii) $\begin{vmatrix} 25 & 30 & 35 \\ 15 & 35 & 40 \\ 10 & 25 & 45 \end{vmatrix}$ (iv) $\begin{vmatrix} 8 & 16 & 24 \\ 24 & 15 & 23 \\ 40 & 1 & 1 \end{vmatrix}$

(v) $\begin{vmatrix} -3 & -4 & 2 \\ 1 & 0 & -2 \\ -5 & -6 & 7 \end{vmatrix}$ (vi) $\begin{vmatrix} 781 & 397 & 542 \\ 115 & 201 & 136 \\ 551 & -5 & 270 \end{vmatrix}$

2. (*a*) If

$$\Delta = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}$$

and A, B, . . . denote the co-factors of $a, b, \ldots$

Evaluate

(i) $a\mathrm{A} + e\mathrm{E} + i\mathrm{I}$.
(ii) $a\mathrm{B} + d\mathrm{E} + g\mathrm{H}$.

(*b*) Find the cofactors of (i) p (ii) q in each of the following

$$\begin{vmatrix} l & m & n \\ p & q & r \\ s & t & u \end{vmatrix}, \begin{vmatrix} a & 0 & q \\ b & p & 1 \\ c & 0 & 1 \end{vmatrix},$$

$$\begin{vmatrix} q & p \\ a & b \end{vmatrix}.$$

3. Find the integral root of

$$\begin{vmatrix} 2-\lambda & 2 & 3 \\ 1 & 6-\lambda & 7 \\ 2 & 7 & 11-\lambda \end{vmatrix} = 0$$

and approximate to the other two roots.

4. Give reasons from the list of properties of a determinant on page 2.4 that enable the following steps to be completed

Reason

$$\begin{vmatrix} 0 & 1 & x \\ x & x & 1 \\ x^2-x & x^2-1 & x^2-1 \end{vmatrix} = (x-1)\begin{vmatrix} 0 & 1 & x \\ x & x & 1 \\ x & x+1 & x+1 \end{vmatrix}$$

$$(x-1)\begin{vmatrix} 0 & 1 & x \\ x & x & 1 \\ x & x+1 & x+1 \end{vmatrix} = (x-1)\begin{vmatrix} 0 & 1 & x \\ x & 0 & 1 \\ x & 1 & x+1 \end{vmatrix}$$

$$(x-1)\begin{vmatrix} 0 & 1 & x \\ x & 0 & 1 \\ x & 1 & x+1 \end{vmatrix} = x(x-1)\begin{vmatrix} 0 & 1 & x \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{vmatrix}$$

Hence evaluate the determinant.

5. Without expanding the determinant in full, show that

$$\begin{vmatrix} p & q & r \\ q+r & r+p & p+q \\ 1 & 1 & 1 \end{vmatrix} = 0$$

6. (*a*) Show that

$$\begin{vmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{vmatrix} = adf$$

(*b*) Hence evaluate without expanding

(i) $\begin{vmatrix} 1 & 0 & 0 \\ 2 & 4 & 0 \\ 3 & 5 & 6 \end{vmatrix}$ (ii) $\begin{vmatrix} 3 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 5 \end{vmatrix}$

(iii) $\begin{vmatrix} 0 & 0 & a \\ 0 & d & b \\ f & e & c \end{vmatrix}$ (iv) $\begin{vmatrix} 0 & 0 & \sqrt{2} \\ 0 & \sqrt{3} & 0 \\ \sqrt{5} & 0 & 0 \end{vmatrix}$

7. If $A = \begin{pmatrix} 2 & 4 \\ 1 & 3 \end{pmatrix}$, evaluate the $n \times n$ determinant

$$\begin{vmatrix} |A| & 0 & 0 & \cdots & 0 \\ 0 & |A| & 0 & \cdots & 0 \\ 0 & 0 & |A| & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & |A| \end{vmatrix}$$

8. If

$$A = \begin{pmatrix} 1 & 2 & 4 \\ -1 & -2 & 3 \\ 7 & 0 & 1 \end{pmatrix}$$

Evaluate

$$|A^2|;\ |5A|;\ 5|A|;\ |\tfrac{1}{3}A'|;\ |A^3A'|$$

9. Factorize

(*a*) $\begin{vmatrix} 1 & p & p^2 \\ 1 & q & q^2 \\ 1 & r & r^2 \end{vmatrix}$ (*b*) $\begin{vmatrix} 1 & p & p^3 \\ 1 & q & q^3 \\ 1 & r & r^3 \end{vmatrix}$

10. Show that

$$\begin{vmatrix} x & y & 1 \\ y & 1 & x \\ 1 & x & y \end{vmatrix} = (x + y + 1)(xy - x^2 - y^2 + x + y - 1)$$

11. (i) Prove that

$$\begin{vmatrix} a & b & c \\ a^2 & b^2 & c^2 \\ bc & ca & ab \end{vmatrix} = \begin{vmatrix} 1 & a^2 & a^3 \\ 1 & b^2 & b^3 \\ 1 & c^2 & c^3 \end{vmatrix}$$

(ii) Factorize

$$\begin{vmatrix} x^2 & y^2 & z^2 \\ (y+z)^2 & (z+x)^2 & (x+y)^2 \\ yz & zx & xy \end{vmatrix}$$

12. Evaluate the determinant

$$\begin{vmatrix} 2 & 1+j & 2 \\ 1-j & -1 & 4-j \\ 2 & 4+j & 3 \end{vmatrix}$$

where $j^2 = -1$. Explain how, without evaluation, it could have been concluded that the value of the determinant was real. (*MEI Pure I*)

13. If ω is a complex cube root of unity, prove that $1 + \omega + \omega^2 = 0$. Find in its simplest form, the value of the determinant,

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega^3 \end{vmatrix}$$

(*MEI Specimen*)

14. Evaluate the determinant

$$\Delta = \begin{vmatrix} 8 & 9 & 10 \\ 4 & 5 & 6 \\ 91 & 92 & 93 \end{vmatrix}$$

Hence or otherwise find the area of the triangle ABC where A is (91, 95), B is (92, 97), and C is (93, 99). (*MEI Specimen*)

15. (i) Prove that

$$\begin{vmatrix} 1 & \dfrac{2x}{x+y} & \dfrac{2}{x+z} \\ \dfrac{2y}{x+y} & 1 & \dfrac{2y}{y+z} \\ \dfrac{2z}{x+z} & \dfrac{2z}{y+z} & 1 \end{vmatrix} = \left[\frac{(y-z)(z-x)(x-y)}{(y+z)(z+x)(x+y)}\right]^2$$

(ii) Prove that the nth-order determinant

$$\begin{vmatrix} x & a & a & a & . & . & . \\ a & x & 0 & 0 & . & . & . \\ a & 0 & x & 0 & . & . & . \\ a & 0 & 0 & x & . & . & . \\ . & . & . & . & . & . & . \end{vmatrix}$$

is equal to $x^n - (n-1)a^2x^{n-2}$. *(O & C)*

16. (i) Solve the following equation for x given that $a^2 \neq 1$.

$$\begin{vmatrix} x^3 & a^2+1 & 1 \\ a^3 & x^2+1 & 1 \\ 1 & x^2+a^2 & 1 \end{vmatrix} = 0$$

(ii) In a determinant of order n, every element in the principal diagonal is equal to x, and every other element is equal to y. Show that the determinant has $(x-y)^{n-1}$ as a factor, and find an expression for the determinant. *(O & C)*

17. (i) Express the determinant

$$\begin{vmatrix} a & b & c & d \\ d & a & b & c \\ c & d & a & b \\ b & c & d & a \end{vmatrix}$$

as a product* of two (real) linear factors and one (real) quadratic factor.

(ii) Express the determinant

$$\begin{vmatrix} x^2+2yz & z^2+2xy & y^2+2zx \\ y^2+2zx & x^2+2yz & z^2+2xy \\ z^2+2xy & y^2+2zx & x^2+2yz \end{vmatrix}$$

as a product of two determinants whose elements are linear in x, y, z. Hence or otherwise express the determinant as the product of linear and quadratic factors. *(O & C)*

18. Factorize the determinants

(i) $\begin{vmatrix} a^2 & (a+1)^2 & (a+2)^2 \\ b^2 & (b+1)^2 & (b+2)^2 \\ c^2 & (c+1)^2 & (c+2)^2 \end{vmatrix}$ (ii) $\begin{vmatrix} x & a & b & c \\ a & x & c & b \\ b & c & x & a \\ c & b & a & x \end{vmatrix}$

(O & C)

---

* The product of two determinants is found in the same way as for matrices.

The inverse matrix

1.11 The $n \times n$ matrix

$$I = \begin{pmatrix} 1 & 0 & 0 & . & . & . & 0 \\ 0 & 1 & 0 & . & . & . & 0 \\ 0 & 0 & 1 & . & . & . & 0 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & . & 1 \end{pmatrix}$$

is known as a unit matrix. Simple examples of unit matrices are

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

It is easily verified that if A is also $n \times n$

$IA = AI = A$, also $AI = A$ if A is $m \times n$; $IA = A$ if A is $n \times p$

The inverse of a square matrix A is that matrix A^{-1}, if it exists, such that $AA^{-1} = A^{-1}A = I$.

The adjoint matrix

1.12 Consider

$$M = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

and the matrix N, whose elements are the cofactors of the elements of $|M|$

$$N = \begin{pmatrix} \mathrm{A} & \mathrm{B} & \mathrm{C} \\ \mathrm{D} & \mathrm{E} & \mathrm{F} \\ \mathrm{G} & \mathrm{H} & \mathrm{I} \end{pmatrix}$$

Now the product

$$MN' = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} \mathrm{A} & \mathrm{D} & \mathrm{G} \\ \mathrm{B} & \mathrm{E} & \mathrm{H} \\ \mathrm{C} & \mathrm{F} & \mathrm{I} \end{pmatrix}$$

$$= \begin{pmatrix} a\mathrm{A} + b\mathrm{B} + c\mathrm{C} & a\mathrm{D} + b\mathrm{E} + c\mathrm{F} & a\mathrm{G} + b\mathrm{H} + c\mathrm{I} \\ d\mathrm{A} + e\mathrm{B} + f\mathrm{C} & d\mathrm{D} + e\mathrm{E} + f\mathrm{F} & d\mathrm{G} + e\mathrm{H} + f\mathrm{I} \\ g\mathrm{A} + h\mathrm{B} + i\mathrm{C} & g\mathrm{D} + h\mathrm{E} + i\mathrm{F} & g\mathrm{H} + h\mathrm{H} + i\mathrm{I} \end{pmatrix}$$

$$= \begin{pmatrix} \Delta & 0 & 0 \\ 0 & \Delta & 0 \\ 0 & 0 & \Delta \end{pmatrix}$$

by the properties of cofactors* where

$$\Delta = |M| = \Delta \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \Delta I$$

Hence $$M \times \left(\frac{1}{\Delta} N'\right) = I$$

and since $$MM^{-1} = I$$

$$M^{-1} = \frac{1}{\Delta} N'$$

N' is known as the adjoint matrix of M and is written '**adj** M.'

Inverses exist only for square matrices since determinants are defined only for square matrices. Other ideas of inverses will be found in other systems.

Inverse of a 3×3 matrix by the adjoint method

1.13 Use is made of the fact that $M^{-1} = \frac{1}{\Delta}\,\mathbf{adj}\,M$. The method is illustrated by a worked example.

Worked Example

To find the inverse of the matrix

$$\begin{pmatrix} 1 & 2 & 4 \\ 0 & 1 & 6 \\ 1 & 3 & 2 \end{pmatrix}$$

The matrix of the cofactors of M is N

$$N = \begin{pmatrix} -16 & 6 & -1 \\ 8 & -2 & -1 \\ 8 & -6 & 1 \end{pmatrix}$$

Transpose N to find **adj** M

$$\mathbf{adj}\,M = N' = \begin{pmatrix} -16 & 8 & 8 \\ 6 & -2 & -6 \\ -1 & -1 & 1 \end{pmatrix}$$

Multiply M by **adj** M

$$M\,.\,\mathbf{adj}\,M = \begin{pmatrix} 1 & 2 & 4 \\ 0 & 11 & 6 \\ 1 & 3 & 2 \end{pmatrix}\begin{pmatrix} -16 & 8 & 8 \\ 6 & -2 & -6 \\ -1 & -1 & 1 \end{pmatrix} = \begin{pmatrix} -8 & 0 & 0 \\ 0 & -8 & 0 \\ 0 & 0 & -8 \end{pmatrix}$$

---

* See p. 10 and Exercise 1.5 Question 2.

Hence $\Delta = |M| = -8$

and
$$M^{-1} = \frac{1}{\Delta}\,\mathbf{adj}\,M = -\frac{1}{8}\begin{pmatrix} -16 & 8 & 8 \\ 6 & -2 & -6 \\ -1 & -1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 2 & -1 & -1 \\ -\frac{3}{4} & \frac{1}{4} & \frac{3}{4} \\ \frac{1}{8} & \frac{1}{8} & -\frac{1}{8} \end{pmatrix}$$

As a check

$$MM^{-1} = \begin{pmatrix} 1 & 2 & 4 \\ 0 & 1 & 6 \\ 1 & 3 & 2 \end{pmatrix}\begin{pmatrix} 2 & -1 & -1 \\ -\frac{3}{4} & \frac{1}{4} & \frac{3}{4} \\ \frac{1}{8} & \frac{1}{8} & -\frac{1}{8} \end{pmatrix}$$

$$= \begin{pmatrix} 2 - \frac{3}{2} + \frac{1}{2} & -1 + \frac{1}{2} + \frac{1}{2} & -1 + \frac{3}{2} - \frac{1}{2} \\ -\frac{3}{4} + \frac{3}{4} & \frac{1}{4} + \frac{3}{4} & \frac{3}{4} - \frac{3}{4} \\ 2 - \frac{9}{4} + \frac{1}{4} & -1 + \frac{3}{4} + \frac{1}{4} & -1 + \frac{9}{4} - \frac{1}{4} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Singular matrices

1.14 A square matrix that has no inverse is said to be *singular*; otherwise it is said to be *non-singular.*

Since $M^{-1} = \frac{1}{\Delta}\,\mathbf{adj}\,M$ it follows that a singular matrix has $\Delta = 0$.

Matrices other than square matrices have no inverses since their determinants are not defined.

Exercise 1.6

1. Use the non-singular matrices

$$A = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & 4 \\ 0 & 1 & 2 \end{pmatrix}, \qquad B = \begin{pmatrix} -1 & 4 & 6 \\ 0 & 3 & -1 \\ 0 & 2 & -1 \end{pmatrix}$$

to demonstrate that $(AB)^{-1} = B^{-1}A^{-1}$.

2. Use the adjoint method to find the inverse (if it exists) of each of the following matrices:

(*a*) $\begin{pmatrix} 5 & 4 \\ 3 & 5 \end{pmatrix}$ (*b*) $\begin{pmatrix} 2 & 4 & -2 \\ 6 & 10 & 4 \\ 4 & 6 & 2 \end{pmatrix}$

(*c*) $\begin{pmatrix} 9 & 15 & 18 \\ 12 & 24 & 21 \\ 21 & 42 & 39 \end{pmatrix}$ (*d*) $\begin{pmatrix} 8 & 3 & 5 \\ 4 & 11 & 7 \\ 3 & 9 & 6 \end{pmatrix}$

(*e*) $\begin{pmatrix} 1 & 3 & 2 \\ 1 & 1 & 1 \\ 2 & -3 & -1 \end{pmatrix}$ (*f*) $\begin{pmatrix} 5 & 6 & 8 \\ 7 & 1 & 4 \\ 3 & 2 & 9 \end{pmatrix}$

3. Given a matrix A, the unit matrix I and $B = I - A$, it may be assumed under certain conditions that

$$A^{-1} = (I - B)^{-1} = I + B + B^2 + \dots$$

This identity leads to an iterative process for calculating the inverse of a matrix A. By means of a flow chart show how the partial sums

$$C_n = I + \sum_1^n B^k$$

might be calculated in as few stages as possible. If

$$A = \begin{pmatrix} \frac{1}{2} & 1 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & -1 & \frac{3}{4} \end{pmatrix}$$

evaluate C_1 and C_2, and also the difference matrix $AC_2 - I$.

(*MEI Specimen*)

4. The non-singular matrix B has the property $BB' = B'B$, where B' is the transpose of B. Prove that $B'B^{-1} = B^{-1}B'$. Prove also that if $C = B^{-1}B'$, then CC' is the identity matrix. Find B', BB', B^{-1} and C when

$$B = \begin{pmatrix} 2 & 2 & 1 \\ -2 & 1 & 2 \\ 1 & -2 & 2 \end{pmatrix}$$

(*MEI Pure I*)

2
Elementary Matrices: Partitioned Matrices

2.1 To find the inverse of matrices of order 4 × 4 and higher, and especially when a computer is used, one of the simplest and quickest methods is with elementary matrices. Merely as a demonstration, the method is investigated for 2 × 2 matrices, but in this case it is long and clumsy.

Exercise 2.1

1. Let

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad \text{and let} \qquad E_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

$$E_2 = \begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix}, \qquad E_3 = \begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix},$$

$$E_4 = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}, \qquad E_5 = \begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix},$$

where $k \in C$, the set of complex numbers.
Evaluate the products

$$(a)\ E_1A, \quad (b)\ E_2A, \quad (c)\ E_3A, \quad (d)\ E_4A, \quad (e)\ E_5A$$

What happens if the order of multiplication is reversed?

2. If

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} -12 & 16 \\ 18 & 20 \end{pmatrix} \quad \text{and} \quad C = \begin{pmatrix} 6 & 9 \\ 3 & 12 \end{pmatrix}$$

use the results of question 1 to find (*a*) LA, (*b*) MB, (*c*) NC in the cases:

(i) $L = M = N = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

(ii) $L = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$, $M = \begin{pmatrix} -\frac{1}{4} & 0 \\ 0 & 1 \end{pmatrix}$, $N = \begin{pmatrix} \frac{1}{3} & 0 \\ 0 & 1 \end{pmatrix}$

(iii) $L = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$, $M = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$, $N = \begin{pmatrix} 1 & 0 \\ 0 & -\frac{1}{3} \end{pmatrix}$

(iv) $L = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $M = \begin{pmatrix} 1 & -\frac{1}{2} \\ 0 & 1 \end{pmatrix}$, $N = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$

(v) $L = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, $M = \begin{pmatrix} 1 & 0 \\ \frac{1}{3} & 1 \end{pmatrix}$, $N = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}$

The 2×2 elementary matrices

2.2 The six matrices,

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

together with E_1, E_2, E_3, E_4, E_5, defined as in question 1 above, are called the 2 × 2 elementary matrices.

If A is the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then

(i) $IA = A$.
(ii) E_1A interchanges the rows of A.
(iii) E_2A multiplies the top row of A by the number k.
(iv) E_3A multiplies the bottom row of A by the number k.
(v) E_4A adds k times the bottom row of A onto the top.
(vi) E_5A adds k times the top row of A onto the bottom.

Similar results are true for the columns of A if the order of multiplication is reversed, but throughout this chapter only row operations will be considered.

To find the inverse of a 2×2 matrix using elementary matrices

2.3 By multiplying in front by a chain of elementary matrices, the given matrix is reduced to the unit matrix. The method is illustrated by a worked example.

Worked Example

Find the inverse of $\begin{pmatrix} 2 & 3 \\ 6 & 13 \end{pmatrix}$

$$\begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix}\begin{pmatrix} 2 & 3 \\ 6 & 13 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 0 & 4 \end{pmatrix}$$

This operation was to produce a zero in the bottom left-hand element.

$$\swarrow$$

$$\begin{pmatrix} 1 & -\frac{3}{4} \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 2 & 3 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$$

This operation was to produce a zero in the top right-hand element.

$$\swarrow$$

$$\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$$

This operation was to produce a 1 in the top left-hand element.

$$\swarrow$$

$$\begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{4} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

This reduces the original matrix to the unit matrix.

Note in full what has been performed:

$$\begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{4} \end{pmatrix}\left\{\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{pmatrix}\left[\begin{pmatrix} 1 & -\frac{3}{4} \\ 0 & 1 \end{pmatrix}\left\{\begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix}\begin{pmatrix} 2 & 3 \\ 6 & 13 \end{pmatrix}\right\}\right]\right\} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

But matrix multiplication is associative. Hence the brackets may be relocated.

$$\left[\begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{4} \end{pmatrix}\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & -\frac{3}{4} \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix}\right]\begin{pmatrix} 2 & 3 \\ 6 & 13 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

But the inverse A^{-1} of a matrix A is that matrix such that $A^{-1}A = I$. Hence if

$$A = \begin{pmatrix} 2 & 3 \\ 6 & 13 \end{pmatrix}$$

$$A^{-1} = \left[\begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{4} \end{pmatrix}\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & -\frac{3}{4} \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix}\right]$$

$$= \begin{pmatrix} \frac{13}{8} & -\frac{3}{8} \\ -\frac{3}{4} & \frac{1}{4} \end{pmatrix}$$

As a check:

$$\begin{pmatrix} \frac{13}{8} & -\frac{3}{8} \\ -\frac{3}{4} & \frac{1}{4} \end{pmatrix}\begin{pmatrix} 2 & 3 \\ 6 & 13 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Exercise 2.2

Form a chain of elementary matrices to derive the inverse of the following matrices:

1. $\begin{pmatrix} 3 & 4 \\ 2 & 3 \end{pmatrix}$
2. $\begin{pmatrix} -1 & -3 \\ -1 & -2 \end{pmatrix}$
3. $\begin{pmatrix} 12 & 13 \\ 10 & 11 \end{pmatrix}$
4. $\begin{pmatrix} a & b \\ b & a \end{pmatrix}$ given $a^2 \neq b^2$
5. (i) Prove that

$$\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} c & d \\ a & b \end{pmatrix}$$

Deduce that the rows of a 2 × 2 matrix can be interchanged by operations which add multiples of one row of the matrix to the other, together with operations changing the sign of a row.

Find a 2 × 2 matrix X such that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} X = \begin{pmatrix} b & a \\ d & c \end{pmatrix}$$

(ii) Do 2 × 2 matrices, A, B, with integer entries, exist such that

(*a*) $AB = O, BA \neq O$?
(*b*) $AB = BA = O, A \neq O, B \neq O, A \neq B$?
(*c*) $AB = BA \neq O, A, B \neq I, A \neq B$?

In each case, if your answer is 'yes', justify it by giving examples of suitable A, B. (*MEI Pure II*)

Better method of finding the inverse of 2×2 matrix

2.4 The preceding method for finding the inverse of a 2 × 2 matrix can be improved upon. In the worked example it was noted that

$$\left[\begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{4} \end{pmatrix}\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & -\frac{3}{4} \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix}\right]$$

was the inverse of

$$\begin{pmatrix} 2 & 3 \\ 6 & 13 \end{pmatrix}$$

Hence

$$\left[\begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{4} \end{pmatrix}\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & -\frac{3}{4} \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix}\right]\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

is also the inverse. But since multiplication is associative, relocating the brackets, the inverse is

$$\begin{pmatrix}1 & 0\\0 & \frac{1}{4}\end{pmatrix}\left\{\begin{pmatrix}\frac{1}{2} & 0\\0 & 1\end{pmatrix}\left[\begin{pmatrix}1 & -\frac{3}{4}\\0 & 1\end{pmatrix}\left\{\begin{pmatrix}1 & 0\\-3 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\right\}\right]\right\}$$

Compare this with the multiplication that was applied to the original matrix to reduce it to the unit matrix.

$$\begin{pmatrix}1 & 0\\0 & \frac{1}{4}\end{pmatrix}\left\{\begin{pmatrix}\frac{1}{2} & 0\\0 & 1\end{pmatrix}\left[\begin{pmatrix}1 & -\frac{3}{4}\\0 & 1\end{pmatrix}\left\{\begin{pmatrix}1 & 0\\-3 & 1\end{pmatrix}\begin{pmatrix}2 & 3\\6 & 13\end{pmatrix}\right\}\right]\right\}$$

This means that the inverse of a matrix can be obtained by finding the elementary row operations that reduce it to the unit matrix; and simultaneously performing these operations on the unit matrix. There is no need to find the elementary matrices used; but merely to perform the row operations that pre-multiplication by them would achieve.

2.5 For example, to find the inverse of

$$\begin{pmatrix}2 & 3\\6 & 13\end{pmatrix}$$

Proceeding with the same row operations as before

Unit matrix $\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}$ *Original matrix* $\begin{pmatrix}2 & 3\\6 & 13\end{pmatrix}$

take 3 times the top row from the bottom, $R_2 - 3R_1$

$$\begin{pmatrix}2 & 3\\0 & 4\end{pmatrix}$$

This operation is simultaneously performed on the unit matrix

$$\begin{pmatrix}1 & 0\\-3 & 1\end{pmatrix} \quad \Big|$$

$$\begin{pmatrix}\frac{13}{4} & -\frac{3}{4}\\-3 & 1\end{pmatrix} \quad \Big| \quad \begin{pmatrix}2 & 0\\0 & 4\end{pmatrix}$$

Now take $\frac{3}{4}$ × Row 2 from Row 1 ($R_1 - \frac{3}{4}R_2$)

$$\begin{pmatrix}\frac{13}{8} & -\frac{3}{8}\\-3 & 1\end{pmatrix} \quad \Big| \quad \begin{pmatrix}1 & 0\\0 & 4\end{pmatrix}$$

Take $\frac{1}{2}$ of Row 1 ($\frac{1}{2}R_1$)

$$\begin{pmatrix}\frac{13}{8} & -\frac{3}{8}\\-\frac{3}{4} & \frac{1}{4}\end{pmatrix} \quad \Big| \quad \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}$$

Take $\frac{1}{4}$ of Row 2 ($\frac{1}{4}R_2$)

Now $\begin{pmatrix} 2 & 3 \\ 6 & 13 \end{pmatrix}$ has been reduced to $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

and simultaneously, using the same row operations,

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ becomes } \begin{pmatrix} \frac{13}{8} & -\frac{3}{8} \\ -\frac{3}{4} & \frac{1}{4} \end{pmatrix}$$

Hence, using the idea in **2.4**,

$$\begin{pmatrix} \frac{13}{8} & -\frac{3}{8} \\ -\frac{3}{4} & \frac{1}{4} \end{pmatrix} \text{ is the inverse of } \begin{pmatrix} 2 & 3 \\ 6 & 13 \end{pmatrix}$$

Exercise 2.3

Use the method of row operations above to find the inverse of the following matrices:

1. $\begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix}$
2. $\begin{pmatrix} 8 & 3 \\ 5 & 2 \end{pmatrix}$
3. $\begin{pmatrix} 9 & 10 \\ 10 & 11 \end{pmatrix}$
4. $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$
5. $\begin{pmatrix} -5 & -6 \\ -3 & -5 \end{pmatrix}$
6. $\begin{pmatrix} p & q \\ r & s \end{pmatrix}$ given $ps - qr \neq 0$

The 3×3 elementary matrices

2.6 The ideas of 2 × 2 elementary matrices carry over for 3 × 3 matrices.

Exercise 2.4

1. If

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

where $k \in C$, the set of complex numbers and E_n is given as below, find E_nA in each case if

(*a*) $E_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (*b*) $E_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ (*c*) $E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$

(*d*) $E_4 = \begin{pmatrix} k & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (*e*) $E_5 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (*f*) $E_6 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & k \end{pmatrix}$

(*g*) $$E_7 = \begin{pmatrix} 1 & 0 & k \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$ (*h*) $$E_8 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & k \\ 0 & 0 & 1 \end{pmatrix}$$ (*i*) $$E_9 = \begin{pmatrix} 1 & k & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

(*j*) $$E_{10} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & k & 1 \end{pmatrix}$$ (*k*) $$E_{11} = \begin{pmatrix} 1 & 0 & 0 \\ k & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$ (*l*) $$E_{12} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ k & 0 & 1 \end{pmatrix}$$

2. Use the results of question 1 to compute the following products. If

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

find EA when

(i) $$E = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$ (ii) $$E = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

(iii) $$E = \begin{pmatrix} 1 & 0 & \frac{1}{3} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$ (iv) $$E = \begin{pmatrix} 1 & 0 & 0 \\ 10 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

(v) $$E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$ (vi) $$E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

2.7 The thirteen matrices,

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

together with $E_1, E_2, \ldots, E_{12}$, defined as in question 1 above, are called the 3 × 3 elementary matrices. If

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

then

(i) $IA = A$.
(ii) E_1A, E_2A, E_3A each interchange two rows of A.
(iii) E_4A, E_5A, E_6A each multiply one row of A by a scalar k.
(iv) $E_7A, E_8A, \ldots, E_{12}A$ each add k times one row of A to another row.

Similar results are shown to be true for the columns of A by reversing the order of multiplication. Here only row operations are considered.

Echelon form

2.8 A matrix which has only zeros below its *leading diagonal* is said to be in *echelon form.*

An example of a 3 × 3 matrix which has this form is

$$\begin{pmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{pmatrix}$$

In this case the leading diagonal is

$$\begin{matrix} a & & \\ & d & \\ & & f \end{matrix}$$

2.9 A 3 × 3 matrix can be reduced to echelon form by pre-multiplication by a chain of elementary matrices. The method is shown by the following example.

Worked Example

Using a chain of elementary matrices reduce

$$\begin{pmatrix} 3 & 0 & 3 \\ 4 & 2 & 6 \\ 12 & 18 & 6 \end{pmatrix}$$

to echelon form.

$$\begin{pmatrix} 3 & 0 & 3 \\ 4 & 2 & 6 \\ 12 & 18 & 6 \end{pmatrix}$$

$$\swarrow$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{pmatrix}\begin{pmatrix} 3 & 0 & 3 \\ 4 & 2 & 6 \\ 12 & 18 & 6 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 3 \\ 4 & 2 & 6 \\ 0 & 18 & -6 \end{pmatrix}$$

This subtracts 4 times the top row from the bottom $R_3 - 4R_1$

$$\swarrow$$

$$\begin{pmatrix} 1 & 0 & 0 \\ -\frac{4}{3} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 3 & 0 & 3 \\ 4 & 2 & 6 \\ 0 & 18 & -6 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 3 \\ 0 & 2 & 2 \\ 0 & 18 & -6 \end{pmatrix} \qquad R_2 - \tfrac{4}{3}R_1$$

$$\swarrow$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -9 & 1 \end{pmatrix}\begin{pmatrix} 3 & 0 & 3 \\ 0 & 2 & 2 \\ 0 & 18 & -6 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 3 \\ 0 & 2 & 2 \\ 0 & 0 & -24 \end{pmatrix} \qquad R_3 - 9R_2$$

$$\swarrow$$

$$\begin{pmatrix} 3 & 0 & 3 \\ 0 & 2 & 2 \\ 0 & 0 & -24 \end{pmatrix}$$ is now in echelon form.

To find the inverse of a 3×3 matrix using elementary matrices

2.10 The improved method for 2 × 2 matrices is used again. By multiplying a non-singular 3 × 3 matrix by a chain of elementary matrices the matrix can be reduced to the unit matrix.

If M is a 3 × 3 matrix and

$$E_n \ldots E_b E_a M = I$$

where $E_n, \ldots, E_a$ are 3 × 3 elementary matrices then

$$(E_n \ldots E_b E_a) = M^{-1}*$$

This means that

$$(E_n \ldots E_a E_b) I = M^{-1}$$

Hence if M is reduced to the unit matrix using a chain of elementary row operations, then the same operations applied to the unit matrix, simultaneously reduce it to M^{-1}.

Worked Example

To find the inverse of

$$\begin{pmatrix} 3 & 0 & 3 \\ 4 & 2 & 6 \\ 12 & 18 & 6 \end{pmatrix}$$

First it is reduced to echelon form

$$\begin{array}{c:c|c}
\begin{array}{c}\textit{Unit}\\ \textit{Matrix}\end{array} & \begin{array}{c}\textit{Original}\\ \textit{Matrix}\end{array} & \begin{array}{c}\textit{Elementary}\\ \textit{Row Operation}\end{array} \\
\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} & \begin{pmatrix} 3 & 0 & 3 \\ 4 & 2 & 6 \\ 12 & 18 & 6 \end{pmatrix} & \\
 & & R_3 - 4R_1 \\
\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{pmatrix} & \begin{pmatrix} 3 & 0 & 3 \\ 4 & 2 & 6 \\ 0 & 18 & -6 \end{pmatrix} & \\
 & & R_2 - \frac{4}{3}R_1 \\
\begin{pmatrix} 1 & 0 & 0 \\ -\frac{4}{3} & 1 & 0 \\ -4 & 0 & 1 \end{pmatrix} & \begin{pmatrix} 3 & 0 & 3 \\ 0 & 2 & 2 \\ 0 & 18 & -6 \end{pmatrix} & \\
 & & R_3 - 9R_2 \\
\begin{pmatrix} 1 & 0 & 0 \\ -\frac{4}{3} & 1 & 0 \\ 8 & -9 & 1 \end{pmatrix} & \begin{pmatrix} 3 & 0 & 3 \\ 0 & 2 & 2 \\ 0 & 0 & -24 \end{pmatrix} &
\end{array}$$

---

* The product $E_n \ldots E_b E_a$ is unambiguous since matrix multiplication is associative.

The original matrix is now in echelon form

$$-\tfrac{1}{24}R_3$$

$$\begin{pmatrix} 1 & 0 & 0 \\ -\frac{4}{3} & 1 & 0 \\ -\frac{1}{3} & \frac{3}{8} & -\frac{1}{24} \end{pmatrix} \Bigg| \begin{pmatrix} 3 & 0 & 3 \\ 0 & 2 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

$$R_1 - 3R_3$$

$$\begin{pmatrix} 2 & -\frac{9}{8} & \frac{1}{8} \\ -\frac{4}{3} & 1 & 0 \\ -\frac{1}{3} & \frac{3}{8} & -\frac{1}{24} \end{pmatrix} \Bigg| \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

$$R_2 - 2R_3$$

$$\begin{pmatrix} 2 & -\frac{9}{8} & \frac{1}{8} \\ -\frac{2}{3} & \frac{1}{4} & \frac{1}{12} \\ -\frac{1}{3} & \frac{3}{8} & -\frac{1}{24} \end{pmatrix} \Bigg| \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

There are now zeros everywhere except down the leading diagonal

$$\tfrac{1}{3}R_1$$

$$\begin{pmatrix} \frac{2}{3} & -\frac{3}{8} & \frac{1}{24} \\ -\frac{2}{3} & \frac{1}{4} & \frac{1}{12} \\ -\frac{1}{3} & \frac{3}{8} & -\frac{1}{24} \end{pmatrix} \Bigg| \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\tfrac{1}{2}R_2$$

$$\begin{pmatrix} \frac{2}{3} & -\frac{3}{8} & \frac{1}{24} \\ -\frac{1}{3} & \frac{1}{8} & \frac{1}{24} \\ -\frac{1}{3} & \frac{3}{8} & -\frac{1}{24} \end{pmatrix} \Bigg| \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The original matrix has been reduced to the unit matrix by a series of row operations, and simultaneously the unit matrix has been transformed into the inverse.

Notice that it is unnecessary to actually find the elementary matrices used.

As a check

$$\begin{pmatrix} \frac{2}{3} & -\frac{3}{8} & \frac{1}{24} \\ -\frac{1}{3} & \frac{1}{8} & \frac{1}{24} \\ -\frac{1}{3} & \frac{3}{8} & -\frac{1}{24} \end{pmatrix} \begin{pmatrix} 3 & 0 & 3 \\ 4 & 2 & 6 \\ 12 & 18 & 6 \end{pmatrix}$$

$$= \begin{pmatrix} 2 - \frac{3}{2} + \frac{1}{2} & -\frac{3}{4} + \frac{3}{4} & 2 - \frac{9}{4} + \frac{1}{4} \\ -1 + \frac{1}{2} + \frac{1}{2} & \frac{1}{4} + \frac{3}{4} & -1 + \frac{3}{4} + \frac{1}{4} \\ -1 + \frac{3}{2} - \frac{1}{2} & \frac{3}{4} - \frac{3}{4} & -1 + \frac{9}{4} - \frac{1}{4} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Exercise 2.5

Use elementary matrices to find the inverse of each of the following matrices.

1. $\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{pmatrix}$

2. $\begin{pmatrix}k & 0 & 0\\ 0 & l & 0\\ 0 & 0 & m\end{pmatrix}$

3. $\begin{pmatrix}1 & k & l\\ 0 & 1 & m\\ 0 & 0 & 1\end{pmatrix}$

4. $\begin{pmatrix}1 & 0 & 0\\ 19 & -1 & -5\\ 3 & 0 & -1\end{pmatrix}$

5. $\begin{pmatrix}a & 0 & 0\\ f & b & 0\\ e & d & c\end{pmatrix}$

6. $\begin{pmatrix}13 & -4 & -2\\ -2 & 1 & 0\\ -2 & 0 & 1\end{pmatrix}$

7. $\begin{pmatrix}1 & -2 & 0\\ -14 & 38 & -9\\ 0 & -1 & 1\end{pmatrix}$

8. $\begin{pmatrix}-1 & 2 & -3\\ 4 & -5 & 6\\ -7 & 8 & -9\end{pmatrix}$

The case of a singular matrix

2.11 Elementary row operations are used to try to find the inverse of

$$\begin{pmatrix}6 & 7 & 5\\ 3 & 4 & 8\\ 9 & 11 & 13\end{pmatrix}$$

Take $R_3 - \frac{3}{2}R_1$

$$\begin{pmatrix}6 & 7 & 5\\ 3 & 4 & 8\\ 0 & \frac{1}{2} & \frac{11}{2}\end{pmatrix}$$

$R_2 - \frac{1}{2}R_1$

$$\begin{pmatrix}6 & 7 & 5\\ 0 & \frac{1}{2} & \frac{11}{2}\\ 0 & \frac{1}{2} & \frac{11}{2}\end{pmatrix}$$

$R_3 - R_1$

$$\begin{pmatrix}6 & 7 & 5\\ 0 & \frac{1}{2} & \frac{11}{2}\\ 0 & 0 & 0\end{pmatrix}$$

Now the matrix is in echelon form but there is a zero in the leading diagonal. It is now impossible to use elementary row operations to reduce the matrix to the form

$$\begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}$$

where none of a, b, c are zero. Hence it is impossible to reduce the matrix to the form

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

It follows that the matrix

$$\begin{pmatrix} 6 & 7 & 5 \\ 3 & 4 & 8 \\ 9 & 11 & 13 \end{pmatrix}$$

is singular, and this can easily be verified by considering the determinant.

Exercise 2.6

Try using elementary row operations to reduce each of the following singular matrices to the unit matrix.

1. $\begin{pmatrix} 2 & 4 & 3 \\ 1 & 1 & 1 \\ 5 & 7 & 6 \end{pmatrix}$

2. $\begin{pmatrix} 3 & 5 & 9 \\ 2 & 4 & 8 \\ 1 & 3 & 7 \end{pmatrix}$

3. $\begin{pmatrix} 5 & 10 & 9 \\ 8 & 9 & 13 \\ 5 & 5 & 8 \end{pmatrix}$

4. $\begin{pmatrix} a & b & c \\ 0 & 0 & d \\ 0 & 0 & e \end{pmatrix}$, $a \neq 0, e \neq 0$.

PARTITIONED MATRICES

2.12 Any matrix may be partitioned by a series of horizontal or vertical straight lines into a set of submatrices. For example,

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 6 & 6 \\ 7 & 8 & 9 \\ 10 & 11 & 12 \end{pmatrix}$$

may be partitioned as

$$\left(\begin{array}{cc:c} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ \hdashline 10 & 11 & 12 \end{array}\right)$$

or

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

where

$$A = \begin{pmatrix} 1 & 2 \\ 4 & 5 \\ 7 & 8 \end{pmatrix} \qquad B = \begin{pmatrix} 3 \\ 6 \\ 9 \end{pmatrix}$$

$$C = (10 \quad 11) \qquad D = (12)$$

Exercise 2.7

Show that

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} + \begin{pmatrix} 11 & 8 & 5 \\ 10 & 7 & 4 \\ 9 & 6 & 3 \end{pmatrix} = \begin{pmatrix} A+E & B+F \\ C+G & D+H \end{pmatrix}$$

where (i)

$$A = \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix}, \quad B = \begin{pmatrix} 3 \\ 6 \end{pmatrix}, \qquad C = (7 \quad 8), \qquad D = (9)$$

$$E = \begin{pmatrix} 11 & 8 \\ 10 & 7 \end{pmatrix}, \quad F = \begin{pmatrix} 5 \\ 4 \end{pmatrix}, \qquad G = (9 \quad 6), \qquad H = (3)$$

and (ii)

$$A = (1), \qquad B = (2 \quad 3), \qquad C = \begin{pmatrix} 4 \\ 7 \end{pmatrix}, \qquad D = \begin{pmatrix} 5 & 6 \\ 8 & 9 \end{pmatrix}$$

$$E = (11), \qquad F = (8 \quad 5), \qquad G = \begin{pmatrix} 10 \\ 9 \end{pmatrix}, \qquad H = \begin{pmatrix} 7 & 4 \\ 6 & 3 \end{pmatrix}$$

Block addition

2.13 The above example shows that the *block addition* of partitioned matrices is possible. It is, however, only possible if the original matrices are of the same order, and are partitioned in the same way.

Consider what happens if an attempt is made to find

$$\begin{pmatrix} A+E & B+F \\ C+G & D+H \end{pmatrix}$$

if

$$A = \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix}, \quad B = \begin{pmatrix} 3 \\ 6 \end{pmatrix}, \quad C = (7 \quad 8), \quad D = (9)$$

and

$$E = (11), \quad F = (8 \quad 5), \quad G = \begin{pmatrix} 10 \\ 9 \end{pmatrix}, \quad H = \begin{pmatrix} 7 & 4 \\ 6 & 3 \end{pmatrix}$$

Exercise 2.8

1. Show that

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} AC \\ \\ BC \end{pmatrix}$$

where

$$A = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}, \quad B = (g \quad h \quad i), \quad C = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

2. Show that

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} AE + BF \\ CE + DF \end{pmatrix}$$

where

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad B = \begin{pmatrix} c \\ f \end{pmatrix}, \quad C = (g \quad h), \quad D = (i)$$

and

$$E = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \qquad F = (3)$$

3. If X is partitioned as

$$\begin{matrix} m_1 \updownarrow \\ m_2 \updownarrow \end{matrix} \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right)$$

$$\underset{n_1}{\longleftrightarrow} \quad \underset{n_2}{\longleftrightarrow}$$

and Y is partitioned as

$$\begin{array}{rc} \begin{array}{r} p_1 \updownarrow \\ p_2 \updownarrow \end{array} & \left(\begin{array}{c|c} E & F \\ \hline G & H \end{array}\right) \\ & \begin{array}{cc} \longleftrightarrow & \longleftrightarrow \\ q_1 & q_2 \end{array} \end{array}$$

where m_i, n_i, p_i, q_i are the integral numbers of rows or columns in each partitioned matrix, find the condition that each of the product matrices exist in the *block product*

$$\begin{pmatrix} AE + BG & AF + BH \\ CE + DG & CF + DH \end{pmatrix}$$

What is the condition that XY exists? Is this condition satisfied by the conditions above?

4. The Pauli spin matrices (associated with Elementary Particle Physics) are

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & -j \\ j & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \text{ where } j^2 = -1.$$

Prove that:

(i) $A^2 = B^2 = C^2 = I.$

(ii) $BC = -CB = jA$
$CA = -AC = jB$
$AB = -BA = jC.$

The Eddington-Dirac matrices are

$$E_1 = \left(\begin{array}{c|c} jA & O \\ \hline O & jA \end{array}\right), \quad E_2 = \left(\begin{array}{c|c} jC & O \\ \hline O & jC \end{array}\right)$$

$$E_3 = \left(\begin{array}{c|c} O & -B \\ \hline B & O \end{array}\right), \quad E_4 = \left(\begin{array}{c|c} -jB & O \\ \hline O & jB \end{array}\right)$$

$$E_5 = \left(\begin{array}{c|c} O & -jB \\ \hline -jB & O \end{array}\right)$$

Show that for all values of p

$$E_p^2 = -I$$

and that for all distinct values of p and q

$$E_pE_q = -E_qE_p$$

Block multiplication

2.14 The preceding exercise leads to the conclusion that two partitioned matrices can be block multiplied in partitioned form provided all of the products of the submatrices exist.

Matrix inversion using partitioned matrices

2.15 The matrix whose inverse is to be found is partitioned so that A and D are square, non-singular matrices.

Now if $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ is the partitioned matrix (where A, B, C and D are matrices) with an inverse $\begin{pmatrix} X & Y \\ Z & W \end{pmatrix}$ then

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} X & Y \\ Z & W \end{pmatrix} = \begin{pmatrix} I & O \\ O & I \end{pmatrix}$$

where I is the unit matrix of the required order, which leads to

$$AX + BZ = I \qquad \text{(i)}$$
$$AY + BW = O \qquad \text{(ii)}$$
$$CX + DZ = O \qquad \text{(iii)}$$
$$CY + DW = I \qquad \text{(iv)}$$

From equations (i) and (iii)

$$AX = I - BZ$$

and

$$-D^{-1}CX = Z$$
$$\Rightarrow AX = I + BD^{-1}CX$$
$$\Rightarrow (A - BD^{-1}C)X = I$$
$$\Rightarrow \quad X = (A - BD^{-1}C)^{-1}$$

and so on, the values X, Y, Z and W being

$$X = (A - BD^{-1}C)^{-1}$$
$$Y = -A^{-1}BW$$
$$Z = -D^{-1}CX$$
$$W = (D - CA^{-1}B)^{-1}$$

Using these equations it is necessary to perform four inversions. Easier equations to work with can be found by combining the equations (i) and (iv) with the four similar equations found by considering

$$\begin{pmatrix} X & Y \\ Z & W \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} I & O \\ O & I \end{pmatrix}$$

The resulting equations require the inversion of only two matrices. These equations are:

$$\begin{aligned} X &= (A - BD^{-1}C)^{-1} \\ Y &= -XBD^{-1} \\ Z &= -D^{-1}CX \\ W &= D^{-1} - ZBD^{-1} \end{aligned}$$

Worked Example

Find the inverse of

$$\begin{pmatrix} -1 & -2 & 1 \\ 7 & 1 & 1 \\ -2 & -2 & 1 \end{pmatrix}$$

This matrix is partitioned as follows (although this is not the only possibility)

$$\left(\begin{array}{c|cc} A & \multicolumn{2}{c}{B} \\ -1 & -2 & 1 \\ \hline 7 & 1 & 1 \\ -2 & -2 & 1 \\ C & \multicolumn{2}{c}{D} \end{array}\right)$$

Then

$$\begin{aligned} X &= \left[(-1) - (-2 \quad 1)\begin{pmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix}\begin{pmatrix} 7 \\ -2 \end{pmatrix}\right]^{-1} \\ &= \left[(-1) - (0 \quad 1)\begin{pmatrix} 7 \\ -2 \end{pmatrix}\right]^{-1} \\ &= [(-1) - (-2)]^{-1} \\ &= (1)^{-1} \\ &= (1) \end{aligned}$$

$$\begin{aligned} Y &= -1(-2 \quad 1)\begin{pmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix} \\ &= -1(0 \quad 1) \\ &= (0 \quad -1) \end{aligned}$$

$$\begin{aligned} Z &= -\begin{pmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix}\begin{pmatrix} 7 \\ -2 \end{pmatrix}(1) \\ &= \begin{pmatrix} -3 \\ -4 \end{pmatrix} \end{aligned}$$

$$W = \begin{pmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix} - \begin{pmatrix} -3 \\ -4 \end{pmatrix} \begin{pmatrix} -2 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix} - \begin{pmatrix} 0 & -3 \\ 0 & -4 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{3} & \frac{8}{3} \\ \frac{2}{3} & \frac{13}{3} \end{pmatrix}$$

Hence the inverse of $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ in this case gives

$$\begin{pmatrix} 1 & 0 & -1 \\ -3 & \frac{1}{3} & \frac{8}{3} \\ -4 & \frac{2}{3} & \frac{13}{3} \end{pmatrix}$$

2.16 This method is clearly applicable for finding the inverses of much larger matrices.

Worked Example

Find the inverse of

$$\begin{array}{c} \begin{array}{cc|cc} \quad A & & \quad B & \end{array} \\ \left(\begin{array}{cc|cc} 3 & -3 & -2 & 4 \\ 0 & 1 & 0 & -1 \\ \hline -2 & 2 & 1 & -2 \\ -2 & 0 & 0 & 1 \end{array}\right) \\ \begin{array}{cc|cc} \quad C & & \quad D & \end{array} \end{array}$$

Partition this matrix as shown. There are four equations to solve as before:

$$X = \left[\begin{pmatrix} 3 & -3 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} -2 & 4 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -2 & 2 \\ -2 & 0 \end{pmatrix}\right]^{-1}$$

$$= \left[\begin{pmatrix} 3 & -3 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 4 & -4 \\ 2 & 0 \end{pmatrix}\right]^{-1}$$

$$= \begin{pmatrix} -1 & 1 \\ -2 & 1 \end{pmatrix}^{-1}$$

$$= \begin{pmatrix} 1 & -1 \\ 2 & -1 \end{pmatrix}$$

$$Y = -\begin{pmatrix}1 & -1\\ 2 & -1\end{pmatrix}\begin{pmatrix}-2 & 4\\ 0 & -1\end{pmatrix}\begin{pmatrix}1 & 2\\ 0 & 1\end{pmatrix}$$

$$= \begin{pmatrix}2 & -1\\ 4 & -1\end{pmatrix}$$

$$Z = -\begin{pmatrix}1 & 2\\ 0 & 1\end{pmatrix}\begin{pmatrix}-2 & 2\\ -2 & 0\end{pmatrix}\begin{pmatrix}1 & -1\\ 2 & -1\end{pmatrix}$$

$$= \begin{pmatrix}2 & -4\\ 2 & -2\end{pmatrix}$$

$$W = \begin{pmatrix}1 & 2\\ 0 & 1\end{pmatrix} - \begin{pmatrix}2 & -4\\ 2 & -2\end{pmatrix}\begin{pmatrix}-2 & 4\\ 0 & -1\end{pmatrix}\begin{pmatrix}1 & 2\\ 0 & 1\end{pmatrix}$$

$$= \begin{pmatrix}1 & 2\\ 0 & 1\end{pmatrix} - \begin{pmatrix}-4 & 4\\ -4 & 2\end{pmatrix}$$

$$= \begin{pmatrix}5 & -2\\ 4 & -1\end{pmatrix}$$

Hence the required inverse is

$$\begin{pmatrix}1 & -1 & 2 & -1\\ 2 & -1 & 4 & -1\\ 2 & -4 & 5 & -2\\ 2 & -2 & 4 & -1\end{pmatrix}$$

Exercise 2.9

Use partitioned matrices to find the inverse of each of the following matrices:

1. $\begin{pmatrix}1 & 2 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}$

2. $\begin{pmatrix}13 & -2 & -2\\ -4 & 1 & 0\\ -2 & 0 & 1\end{pmatrix}$

3. $\begin{pmatrix}-3 & -1 & 2\\ 6 & 4 & -5\\ 9 & -7 & 8\end{pmatrix}$

4. $\begin{pmatrix}7 & 14 & 13\\ 4 & 8 & 7\\ 3 & 5 & 6\end{pmatrix}$

5. $\begin{pmatrix}1 & 1 & 2\\ 1 & 2 & -1\\ 1 & 3 & -3\end{pmatrix}$

6. $\begin{pmatrix}1 & 2 & 3\\ 2 & 3 & 4\\ 3 & 4 & 7\end{pmatrix}$

7. $\begin{pmatrix} -1 & 2 & 4 & -1 \\ -2 & 2 & 4 & -1 \\ -1 & 1 & 2 & -1 \\ -4 & 2 & 5 & -2 \end{pmatrix}$

8. $\begin{pmatrix} 1 & 2 & 2 & 2 \\ 1 & 3 & 2 & 2 \\ 1 & 2 & 3 & 2 \\ 1 & 2 & 2 & 3 \end{pmatrix}$

9. $\begin{pmatrix} 1 & 1 & -\frac{1}{2} & \frac{1}{4} \\ 1 & -1 & \frac{1}{2} & -\frac{1}{4} \\ 1 & 1 & \frac{1}{2} & -\frac{1}{4} \\ 1 & 1 & -\frac{1}{2} & -\frac{1}{4} \end{pmatrix}$

10. $\begin{pmatrix} 4 & 3 & 3 & 3 \\ 3 & 3 & 2 & 2 \\ 3 & 2 & 3 & 2 \\ 3 & 2 & 2 & 3 \end{pmatrix}$

3
Linear Equations

The determinant method (Dr Cramer's rule)

3.1 One solution of simultaneous linear equations is provided by making use of determinants.

From Exercise 1.5, question 2, it follows that if

$$\Delta = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}$$

then the sum of products of the elements of any row with their corresponding cofactors is the value of the determinant, e.g.

$$a\mathrm{A} + d\mathrm{D} + g\mathrm{G} = \Delta$$

but the sum of these elements with cofactors from another row is zero, e.g.

$$a\mathrm{B} + d\mathrm{E} + g\mathrm{H} = 0$$

Now consider the simultaneous equations

$$\begin{cases} ax + by + cz = p & ① \\ dx + ey + fz = q & ② \\ gx + hy + iz = r & ③ \end{cases}$$

where p, q, r are not all zero.

If A, B, C, . . . are the co-factors of $a, b, c, \ldots$ in Δ above, then

$$(\mathrm{A} \times ①) + (\mathrm{D} \times ②) + (\mathrm{G} \times ③)$$

gives

$$(a\mathrm{A} + d\mathrm{D} + g\mathrm{G})x + (b\mathrm{A} + e\mathrm{D} + h\mathrm{G})y + (c\mathrm{A} + f\mathrm{D} + i\mathrm{G})z = p\mathrm{A} + q\mathrm{D} + r\mathrm{G}$$

Now $(a\mathrm{A} + d\mathrm{D} + g\mathrm{G}) = \Delta$, and the coefficients of y and z are both zero. Hence

$$x = \frac{[p\mathrm{A} + q\mathrm{D} + r\mathrm{G}]}{\Delta} \qquad (\text{provided } \Delta \neq 0)$$

$$= \frac{\begin{vmatrix} p & b & c \\ q & e & f \\ r & h & i \end{vmatrix}}{\Delta} = \frac{\mathrm{L}}{\Delta}, \text{ say}$$

Similarly $$y = \frac{\begin{vmatrix} a & p & c \\ d & q & f \\ g & r & i \end{vmatrix}}{\Delta} = \frac{\mathrm{M}}{\Delta}; \qquad z = \frac{\begin{vmatrix} a & b & p \\ d & e & q \\ g & h & r \end{vmatrix}}{\Delta} = \frac{\mathrm{N}}{\Delta}$$

Notice that the column $\begin{pmatrix} p \\ q \\ r \end{pmatrix}$ replaces the column of the coefficients in Δ of the required variable.

The result is easily modified for two equations in two unknowns.

Worked Example

To solve $\left.\begin{aligned} 2x + y &= 4 \\ 3x + 5y &= 13 \end{aligned}\right\}$

$$x = \frac{\begin{vmatrix} 4 & 1 \\ 13 & 5 \end{vmatrix}}{\begin{vmatrix} 2 & 1 \\ 3 & 5 \end{vmatrix}} = \frac{7}{7} = 1; \qquad y = \frac{\begin{vmatrix} 2 & 4 \\ 3 & 13 \end{vmatrix}}{\begin{vmatrix} 2 & 1 \\ 3 & 5 \end{vmatrix}} = \frac{14}{7} = 2.$$

Exercise 3.1

Use the determinant method to solve the simultaneous equations:

1. $\left.\begin{aligned} x + y &= 3 \\ 2x + 3y &= 8 \end{aligned}\right\}$

2. $\left.\begin{aligned} x + y + 2z &= 8 \\ 2x - y + 3z &= 15 \\ 3x + 2y - z &= -3 \end{aligned}\right\}$

3. $\left.\begin{aligned} 2x + 3y - z &= 1 \\ x - 4y + 2z &= -6 \\ 5x - 2y + 2z &= -8 \end{aligned}\right\}$

4. $\left.\begin{aligned} x + y - 3z &= -5 \\ x - 3y + z &= 11 \\ x - y - z &= 3 \end{aligned}\right\}$

Some special cases

3.2 If the above method is employed to find the solution of

$$\left.\begin{aligned} ax + by + cz &= p \\ dz + ey + fz &= q \\ gx + hy + iz &= r \end{aligned}\right\}$$

and it is found that the determinants of the coefficients

$$\Delta = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = 0$$

then a **unique solution** does not exist. In this case no solution of the equations will exist, unless the other determinants L, M, N are also zero. If, for example, it is found that L is zero, any value can be assigned to x, say α, and the problem remaining is to solve the equation for y and z in terms of α. A more detailed analysis will be considered later.

Exercise 3.2

Use the above method to find the solution, if it exists, of the equations:

1. $$\left.\begin{aligned} x+y+z&=r\\ x+2y+z&=s\\ 6x+5y+4z&=t\end{aligned}\right\}\quad \begin{aligned}&\text{(i) when } r=3,\ s=6,\ t=8,\\ &\text{(ii) when } r=2,\ s=6,\ t=8.\end{aligned}$$

2. $$\left.\begin{aligned} 15x+5y-10z&=20\\ -9x-3y+6z&=-12\\ lx+my+nz&=16\end{aligned}\right\}\quad \begin{aligned}&\text{(i) when } l=m=n=1,\\ &\text{(ii) when } l=12,\ m=4,\ n=-8.\end{aligned}$$

MATRIX AND GEOMETRIC METHODS

Two equations in two unknowns

3.3 Consider the equations

$$\begin{cases} ax+by=p & \text{(not both } a \text{ and } b \text{ zero)}\\ cx+dy=q & \text{(not both } c \text{ and } d \text{ zero)}\end{cases}$$

These can be written in matrix form as

$$\begin{pmatrix} a & b\\ c & d\end{pmatrix}\begin{pmatrix} x\\ y\end{pmatrix}=\begin{pmatrix} p\\ q\end{pmatrix}$$

and solved by multiplying both sides of the equation by the inverse of

$$\begin{pmatrix} a & b\\ c & d\end{pmatrix}$$

$$\text{i.e. } \frac{1}{\Delta}\begin{pmatrix} d & -b\\ -c & a\end{pmatrix}\begin{pmatrix} a & b\\ c & d\end{pmatrix}\begin{pmatrix} x\\ y\end{pmatrix}=\frac{1}{\Delta}\begin{pmatrix} d & -b\\ -c & a\end{pmatrix}\begin{pmatrix} p\\ q\end{pmatrix}$$

where $\Delta = ad - bc \neq 0$.

$$\Rightarrow \begin{pmatrix} 1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix} x\\ y\end{pmatrix}=\frac{1}{\Delta}\begin{pmatrix} dp-bq\\ -cp+aq\end{pmatrix}$$

$$\Rightarrow \qquad \begin{pmatrix} x\\ y\end{pmatrix}=\frac{1}{\Delta}\begin{pmatrix} dp-bq\\ aq-cp\end{pmatrix}$$

$$\Rightarrow x=\frac{dp-bq}{ad-bc},\qquad y=\frac{aq-cp}{ad-bc}$$

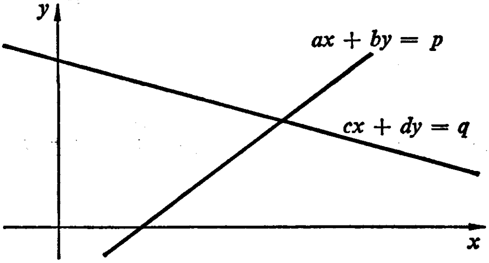
3.4 The solution above is correct in the general case, but it does not give a solution if the determinant of the coefficients, $\Delta = ad - bc = 0$.

To see what happens in this case an appeal is made to geometry.

In general, the equations

$$\left.\begin{aligned} ax + by &= p \\ cx + dy &= q \end{aligned}\right\}$$

can be represented by two intersecting straight lines in a plane.



The coordinates of the point of intersection of the two lines give the solution of the two algebraic equations

$$\left.\begin{aligned} ax + by &= p \\ cx + dy &= q \end{aligned}\right\}$$

Worked Example

Solve the equations $\begin{cases} 2x + y = 4 \\ 3x + 5y = 13 \end{cases}$

The equations can be put in the form

$$\begin{pmatrix} 2 & 1 \\ 3 & 5 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4 \\ 13 \end{pmatrix}$$

Mulitiplying through by the inverse of the matrix of the coefficients

$$\frac{1}{7}\begin{pmatrix} 5 & -1 \\ -3 & 2 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 3 & 5 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{7}\begin{pmatrix} 5 & -1 \\ -3 & 2 \end{pmatrix}\begin{pmatrix} 4 \\ 13 \end{pmatrix}$$

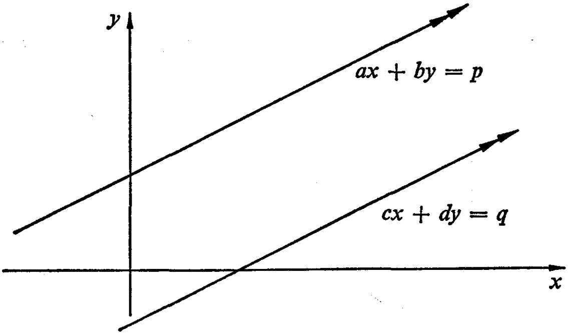
$$\Rightarrow \frac{1}{7}\begin{pmatrix} 7 & 0 \\ 0 & 7 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{7}\begin{pmatrix} 7 \\ 14 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$$\Rightarrow \begin{cases} x = 1 \\ y = 2 \end{cases}$$

Two special cases

3.5 (1) When the two lines are parallel.



In this case the two lines do not intersect, and so there is no solution of the equations

$$\left.\begin{aligned} ax + by &= p \\ cx + dy &= q \end{aligned}\right\}$$

The equations are said to be *inconsistent.*

Since the lines are parallel the equations must be in the form

$$\left.\begin{aligned} ax + by &= p \\ kax + kby &= q \end{aligned}\right\}$$

In this case it is seen that

$$\Delta = \begin{vmatrix} a & b \\ ka & kb \end{vmatrix} = 0$$

Worked Example

Solve the equations

$$\left.\begin{aligned} 6x + 3y &= 7 \\ 4x + 2y &= 1 \end{aligned}\right\}$$

Since

$$\begin{vmatrix} 6 & 3 \\ 4 & 2 \end{vmatrix} = 0$$

the equations cannot be solved as they are inconsistent.

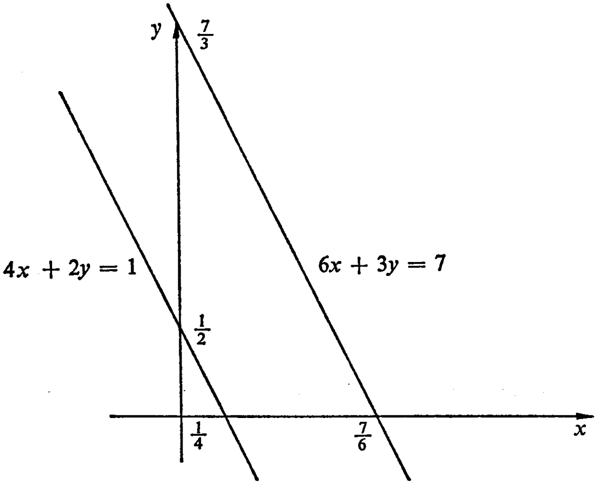
It is seen graphically that the equations represent two parallel lines since they both have a gradient of −2.

Hence there is no solution. It should be noted that $\frac{2}{3}$ times the first equation gives

$$4x + 2y = \tfrac{14}{3}$$

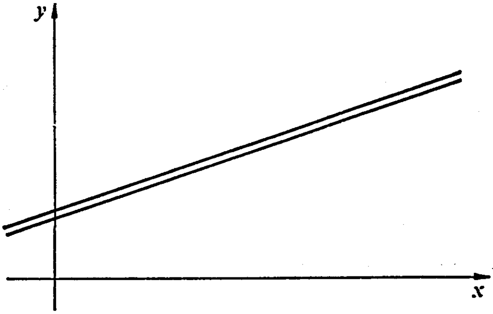
which is inconsistent with the second

$$4x + 2y = 1$$



3.6 (2) When the two lines are coincident. This is a special case of (1). When the equations are in the form

$$\left.\begin{aligned} ax + by &= p \\ kax + kby &= kp \end{aligned}\right\}$$



Again, $\Delta = 0$, but in this case any point on the line will satisfy both the equations. There are an infinite number of solutions.

Worked Example

Solve the equations

$$\left.\begin{aligned} 6x - y &= 5 \\ -12x + 2y &= -10 \end{aligned}\right\}$$

It is seen that

$$\Delta = \begin{vmatrix} 6 & -1 \\ -12 & 2 \end{vmatrix} = 0$$

and so the equations cannot be solved by the matrix method.

It is seen that if the first equation is multiplied through by -2, $6x - y = 5$ becomes $-12x + 2y = -10$, the second equation. One equation is said to be *linearly dependent* upon the other. Hence the two equations are *consistent* and represent the same straight line. There are then an infinity of solutions. If x is given the arbitrary value α then $y = 6\alpha - 5$. The solution is the set of points $(\alpha, 6\alpha - 5)$ for all real values of α.

Exercise 3.3

1. Find the solution set of the following pairs of simultaneous equations:

(*a*) $\left.\begin{aligned} x + y &= 3 \\ x - y &= 1 \end{aligned}\right\}$ (*b*) $\left.\begin{aligned} 5x - 3y &= 10 \\ 3x - 5y &= 15 \end{aligned}\right\}$ (*c*) $\left.\begin{aligned} 27x - 24y &= 6 \\ 18x - 16y &= 9 \end{aligned}\right\}$

(*d*) $\left.\begin{aligned} 27x - 24y &= 9 \\ 18x - 16y &= 6 \end{aligned}\right\}$ (*e*) $\left.\begin{aligned} -2x + y &= 3 \\ x - 2y &= 3 \end{aligned}\right\}$ (*f*) $\left.\begin{aligned} 7x + 8y &= 11 \\ 6x + 7y &= 13 \end{aligned}\right\}$

2. Is (1, 4) an element of the solution set of the equations

$$\left.\begin{aligned} 2x - y &= -2 \\ 2x + y &= 6 \\ x + y &= 5 \end{aligned}\right\} ?$$

Sketch the graph of each equation. Show how it illustrates the answer.

3. Find the solution set of the equations

$$\left.\begin{aligned} x - y &= 1 \\ 2x - y &= 3 \\ 4x - 2y &= 6 \end{aligned}\right\}$$

4. Investigate the conditions under which there is a solution set to the equation

$$\left.\begin{aligned} x + y &= 2 \\ x - y &= 0 \\ ax + by &= c \end{aligned}\right\}$$

Illustrate your answer graphically in three different cases.

5. Show that any equation of the form

$$\lambda(ax + by - p) + \mu(cx + dy - q) = 0 \text{ represents a line}$$

which passes through the point of intersection of the lines

$$\left.\begin{aligned} ax + by &= p \\ cx + dy &= q \end{aligned}\right\}, \text{ where } \lambda, \mu \in R$$

6. Compute the product

$$\begin{pmatrix} -2 - \lambda & 1 - \lambda & \lambda \\ 3 - 2\mu & -2\mu - 1 & 2\mu \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 4 & 6 \end{pmatrix}$$

Use your solution to solve the equations

$$\left.\begin{aligned} x + 2y &= 3 \\ 3x + 4y &= 7 \\ 4x + 6y &= 10 \end{aligned}\right\}$$

7. By considering the possible geometrical configurations, or otherwise, investigate the conditions under which the equations

$$\left.\begin{aligned} ax + by &= p \\ cx + dy &= q \\ ex + fy &= r \end{aligned}\right\}$$

have (*a*) a single solution, (*b*) an infinite number of solutions.

Three equations in two unknowns

3.7 Consider the three equations

$$\left.\begin{aligned} ax + by &= p \\ cx + dy &= q \\ ex + fy &= r \end{aligned}\right\}$$

Solving the last two by the determinant method

$$x = \frac{\begin{vmatrix} q & d \\ r & f \end{vmatrix}}{\begin{vmatrix} c & d \\ e & f \end{vmatrix}}, \qquad y = \frac{\begin{vmatrix} c & q \\ e & r \end{vmatrix}}{\begin{vmatrix} c & d \\ e & f \end{vmatrix}}$$

provided

$$\begin{vmatrix} c & d \\ e & f \end{vmatrix} \neq 0$$

These values of x and y satisfy the first equation if

$$a\begin{vmatrix} q & d \\ r & f \end{vmatrix} + b\begin{vmatrix} c & q \\ e & r \end{vmatrix} = p\begin{vmatrix} c & d \\ e & f \end{vmatrix}$$

$$\Rightarrow a\begin{vmatrix} d & q \\ f & r \end{vmatrix} - b\begin{vmatrix} c & q \\ e & r \end{vmatrix} + p\begin{vmatrix} c & d \\ e & f \end{vmatrix} = 0$$

$$\Rightarrow \begin{vmatrix} a & b & p \\ c & d & q \\ e & f & r \end{vmatrix} = 0$$

If

$$\begin{vmatrix} c & d \\ e & f \end{vmatrix} = 0$$

then the last two equations are inconsistent, unless they are the same equation, and so the given three equations have no solution.

If

$$\begin{vmatrix} c & d \\ e & f \end{vmatrix} = 0$$

and the equations are the same, then either of the last two equations can be solved with the first one.

Exercise 3.4

Use the determinants to find the solution set of the following equations.

1. $\begin{cases} x + y = 5 \\ 2x + 3y = 13 \\ 7x - y = 11 \end{cases}$

2. $\left.\begin{aligned} 2x + 3y &= -1 \\ 3x - 2y &= -5 \\ 7x - 2y &= 9 \end{aligned}\right\}$

3. $\left.\begin{aligned} 11x + 4y &= -38 \\ 2x + 4y &= -20 \\ 3x + 6y &= -30 \end{aligned}\right\}$

4. $\left.\begin{aligned} 32x - 24y &= 16 \\ -8x + 6y &= -4 \\ 3x - \tfrac{9}{4}y &= \tfrac{3}{2} \end{aligned}\right\}$

5. $\left.\begin{aligned} 7x + 3y &= 4 \\ 10x - 2y &= 6 \\ -15x + 3y &= -9 \end{aligned}\right\}$

6. $\left.\begin{aligned} 2x + 3y &= -5 \\ -3x - \tfrac{9}{2}y &= -\tfrac{15}{2} \\ 6x + 18y &= 30 \end{aligned}\right\}$

7. $\left.\begin{aligned} &ax + by = p \\ &cx + dy = q \\ &(a + \lambda c)x + (b + \lambda d)y = r \end{aligned}\right\}$

where

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0$$

in the cases (i) $r \neq p + \lambda q$ (ii) $r = p + \lambda q$

Three equations in three unknowns

3.8 Consider the equations

$$\left.\begin{aligned} ax + by + cz &= p \\ dx + ey + fz &= q \\ gx + hy + iz &= r \end{aligned}\right\}$$

These can be written in the form

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} p \\ q \\ r \end{pmatrix}$$

Writing this as

$$A\mathbf{x} = \mathbf{p}$$

then

$$\begin{aligned} A^{-1}A\mathbf{x} &= A^{-1}\mathbf{p} \\ \Rightarrow I\mathbf{x} &= A^{-1}\mathbf{p} \\ \Rightarrow \mathbf{x} &= A^{-1}\mathbf{p} \end{aligned}$$

i.e.

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = A^{-1}\begin{pmatrix} p \\ q \\ r \end{pmatrix}$$

In general, this will give the unique solution of the three simultaneous equations. The difficulties arise if the matrix A is singular.

To examine the difficulties in this case, i.e. when A is singular

$$\Rightarrow \Delta = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = 0$$

an appeal is again made to geometry.

The equation $ax + by + cz = p$

3.9 Just as the set of ordered pairs (x, y), can be put in one—one correspondence with the points of a plane, so the set of ordered triplets (x, y, z) can be put in one—one correspondence with the points of three-dimensional space.

Just as the equation $ax + by = p$ corresponds to a straight line in the plane, so the equation

$$ax + by + cz = p$$

corresponds to a plane in three-dimensional space.

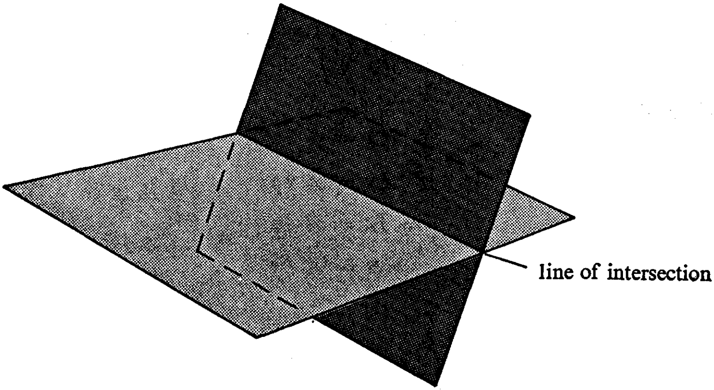
The intersection of three planes in space

3.10 To solve the three algebraic equations

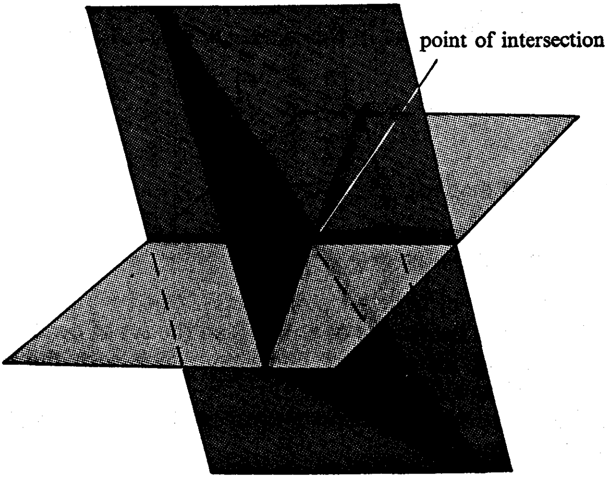
$$\left.\begin{aligned} ax + by + cz &= p \\ dx + ey + fz &= q \\ gx + hy + iz &= r \end{aligned}\right\}$$

the intersection of the three corresponding planes is considered.

In general two planes meet in a straight line



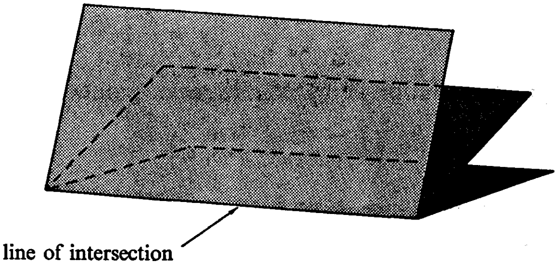
In general, a third plane meets this line of intersection in a single point.



There are three lines of intersection of the planes taken in pairs. These three lines meet in a single point, the point of intersection of the three planes. The coordinates of this point give the solution to the original three simultaneous equations.

The set of coordinates of the point illustrated above is the one given by the matrix solution to the three equations. Several special cases can occur and these are individually examined.

3.11 (*a*) The three planes could intersect like the pages of a book.



In this case the third plane passed through the line of intersection of the other two. The three equations then have the form

$$ax + by + cz = p$$

$$dx + cy + fz = q$$

$$(\lambda a + \mu d)x + (\lambda b + \mu e)y + (\lambda c + \mu f)z = (\lambda p + \mu q)$$

where λ, μ are constants.

Notice that the third equation is a sum of linear multiples of the first two. Hence we say that the third equation is linearly dependent upon the other two and the equations are consistent. There is an infinite set of solutions corresponding to the points of the line of intersection. Notice in this case the determinant of the coefficients.

$$\Delta = \begin{vmatrix} a & b & c \\ d & e & f \\ \lambda a + \mu d & \lambda b + \mu e & \lambda c + \mu f \end{vmatrix} = 0$$

Worked Example

$$\left.\begin{aligned} x + y + z &= 3 \\ 2x - y + 3z &= 4 \\ 4x + y + 5z &= 10 \end{aligned}\right\}$$

It is noted that

$$\begin{vmatrix} 1 & 1 & 1 \\ 2 & -1 & 3 \\ 4 & 1 & 5 \end{vmatrix} = (-5 - 3) - (10 - 12) + (2 + 4) = 0$$

It is also noted that the third equation can be obtained by adding linear multiples of the first two. In fact twice the first equation plus the second produces the third. This means that the third equation is superfluous—it is dependent on the other two.

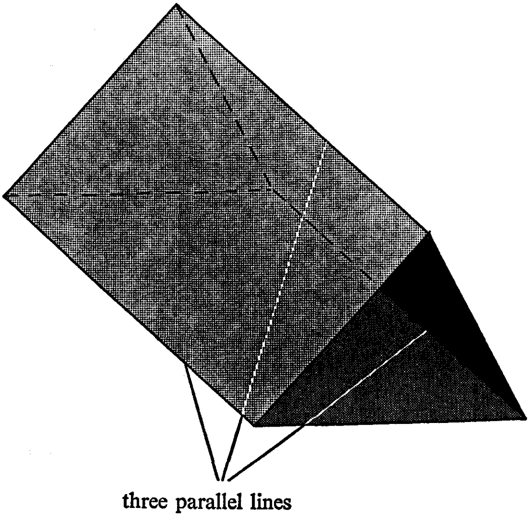
To solve the first two equations, let z have the arbitrary value α, then

$$\left.\begin{aligned} x + y &= 3 - \alpha \\ 2x - y &= 4 - 3\alpha \end{aligned}\right\} \Rightarrow \begin{cases} x = \dfrac{7 - 4\alpha}{3} \\ y = \dfrac{2 + \alpha}{3} \end{cases}$$

Hence the general solution is the infinite set of points

$$\left(\frac{7 - 4\alpha}{3}, \frac{2 + \alpha}{3}, \alpha\right)$$

3.12 (*b*) The three planes could intersect as the rectangular sides of a triangular prism.



In this case there is no common point of intersection of the three planes and hence no solution set. Note that the line of intersection of any two planes is parallel to the third plane.

The three equations in this case will be of the form

$$\left.\begin{aligned} ax + by + cz &= p \\ dx + ey + fz &= q \\ (\lambda a + \mu d)x + (\lambda b + \mu d)y + (\lambda c + \mu f)z &= r \end{aligned}\right\}$$

where $r \neq \lambda p + \mu q$.

Hence we say that the equations are linearly independent and since there is no solution they are inconsistent.

If this last constraint is removed, the planes meet in a line as in (*a*) above. The determinant of the coefficients is as in (*a*). Hence again

$$\Delta = 0$$

Worked Example

Find the solution set of the equations

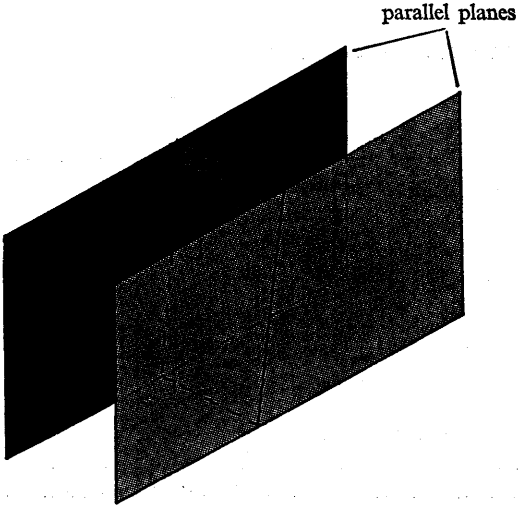
$$\left.\begin{aligned} x + y + z &= 3 \\ 2x - y + 3z &= 4 \\ 4x + y + 5z &= 9 \end{aligned}\right\}$$

It is noted that

$$\Delta = \begin{vmatrix} 1 & 1 & 1 \\ 2 & -1 & 3 \\ 4 & 1 & 5 \end{vmatrix} = 0$$

It is also noted that twice the first equation plus the second gives $4x + y + 5z = 10$, which is inconsistent with the third equation. Hence the solution set is empty.

3.13 (*c*) Two of the planes may be parallel



Again there is no common point of intersection of the three plans, and hence no solution set. The three equations are of the form

$$\begin{cases} ax + by + cz = p \\ kax + kby + kcz = q \\ gx + hy + iz = r \end{cases}$$

where $q \neq kp$.

The determinant of the coefficients is

$$\Delta = \begin{vmatrix} a & b & c \\ ka & kb & kc \\ g & h & i \end{vmatrix} = k \begin{vmatrix} a & b & c \\ a & b & c \\ g & h & i \end{vmatrix} = 0$$

The equations are linearly independent and inconsistent.

Worked Example

Find the solution set of the equations

$$\left.\begin{aligned} 2x - 3y + z &= 4 \\ 4x - 6y + 2z &= 5 \\ x - y - z &= 1 \end{aligned}\right\}$$

The determinant of the coefficients is

$$\Delta = \begin{vmatrix} 2 & -3 & 1 \\ 4 & -6 & 2 \\ 1 & -1 & -1 \end{vmatrix} = 0$$

Note that twice the first equation gives

$$4x - 6y + 2z = 8$$

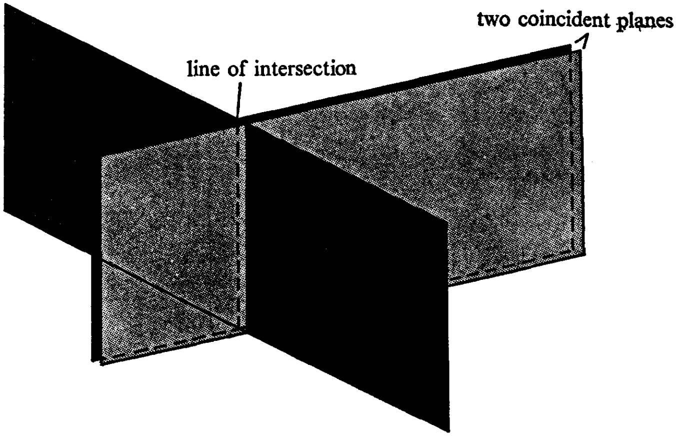
which is inconsistent with the second equation. Hence the corresponding planes are parallel and the solution set is empty.

3.14 (*d*) If in (*c*) the case is considered where $q = kp$ the three equations become

$$\left.\begin{aligned} ax + by + cz &= p \\ kax + kby + kcz &= kp \\ gx + hy + iz &= r \end{aligned}\right\} \text{ with } k \neq 0$$

In this case the second equation is identical with the first (one is linearly dependent upon the other). The determinant of the coefficients is exactly as in (*c*) and is zero. The two corresponding planes are identical.

In general the third plane will intersect these two coincident planes in a straight line, and the coordinates of the set of points of this line will give the solution set of the three equations.



Worked Example

Find the solution set of the equations

$$\left.\begin{aligned} 2x - 3y + z &= 4 \\ 4x - 6y + 2z &= 8 \\ x - y - z &= 1 \end{aligned}\right\}$$

The determinant of the coefficients is again zero as in (*c*) above. The second equation is twice the first, and so the corresponding planes are identical.

To solve the first and third equation, let x take the value α. Hence

$$\left.\begin{aligned} -3y + z &= 4 - 2\alpha \\ -y - z &= 1 - \alpha \end{aligned}\right\}$$

$$\Rightarrow y = \frac{3\alpha - 5}{4},$$

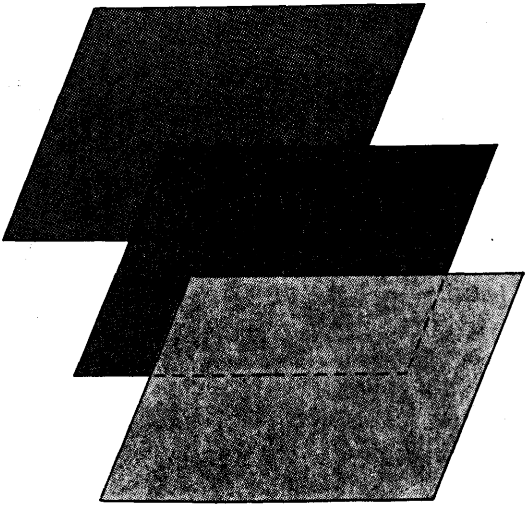
$$z = \frac{7\alpha - 9}{4}$$

The solution is the infinite set

$$\left(\alpha, \frac{3\alpha - 5}{4}, \frac{7\alpha - 9}{4}\right)$$

where $\alpha \in R$.

3.15 (*e*) All three planes can be parallel



The three equations have the form

$$\left.\begin{aligned} ax + by + cz &= p \\ \lambda ax + \lambda by + \lambda cz &= q \\ \mu ax + \mu by + \mu cz &= r \end{aligned}\right\} p \neq \lambda q \neq \mu r; \qquad \lambda, \mu \neq 0$$

Since there is no common point of intersection of the planes, the solution set of the three equations is empty. Again it is noted that

$$\Delta = \begin{vmatrix} a & b & c \\ \lambda a & \lambda b & \lambda c \\ \mu a & \mu b & \mu c \end{vmatrix} = 0$$

$\Rightarrow$ The equations are linearly independent and inconsistent.

Worked Example

Find the solution set of the equations

$$\left.\begin{aligned} x - 2y + 3z &= 4 \\ 2x - 4y + 6z &= 5 \\ -3x + 6y - 9z &= -1 \end{aligned}\right\}$$

It is noted that

$$\Delta = \begin{vmatrix} 1 & -2 & 3 \\ 2 & -4 & 6 \\ -3 & 6 & -9 \end{vmatrix} = 0$$

Also twice the first equation gives

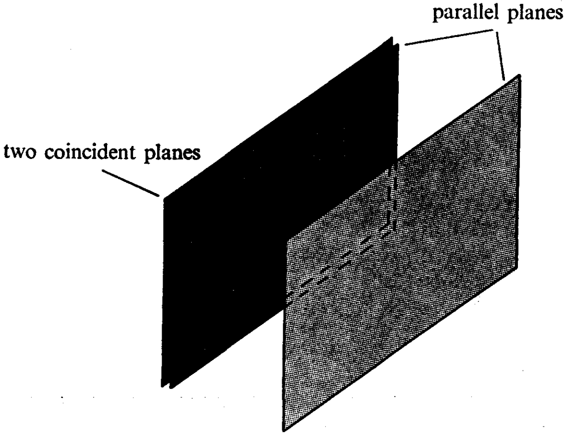
$$2x - 4y + 6z = 8$$

which is inconsistent with the second; minus three times the first gives

$$-3x + 6y - 9z = -12$$

which is inconsistent with the third equation. The three corresponding planes are parallel and the solution set is empty.

3.16 (*f*) In (*e*) two of the planes may be identical and the third parallel to these.



The three equations are as in (*e*) but also $q = \lambda p$. There is no common point of the three planes and hence the solution set is empty. As before $\Delta = 0$.

Worked Example

Find the solution set of the equations

$$\left.\begin{aligned} 2x - y + z &= 3 \\ 6x - 3y + 3z &= 9 \\ 4x - 2y + 2z &= 5 \end{aligned}\right\}$$

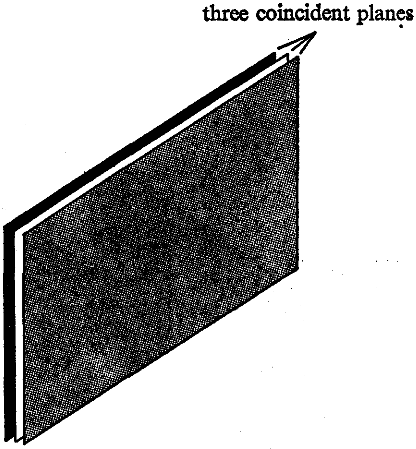
$$\Delta = \begin{vmatrix} 2 & -1 & 1 \\ 6 & -3 & 3 \\ 4 & -2 & 2 \end{vmatrix} = 0$$

Three times the first equation gives the second, and hence the two corresponding planes are identical. Twice the first equation gives

$$4x - 2y + 2z = 6$$

which is inconsistent with the third equation. The first two planes are identical and the third is parallel to these. The solution set is empty.

3.17 (*g*) The equations could be as in (*f*) with the further condition that $r = \mu p$. The three planes would then be coincident, and their corresponding equations identical. The equations are linearly dependent and consistent.



As before $\Delta = 0$.

Since the planes are coincident, the coordinates of any point of the plane would satisfy the three equations.

Worked Example

Find the solution set of the equations

$$\left.\begin{aligned} 3x - y + 2z &= 4 \\ 6x - 2y + 4z &= 8 \\ 9x - 3y + 6z &= 12 \end{aligned}\right\}$$

$$\Delta = \begin{vmatrix} 3 & -1 & 2 \\ 6 & -2 & 4 \\ 9 & -3 & 6 \end{vmatrix} = 0$$

Note that multiplying the first equation by two produces the second, and multiplying the first by three produces the third. The three equations are identical, and hence their corresponding planes are coincident. The solution set is the infinite set of points lying in the plane. Let

$$x = \alpha \quad \text{and} \quad y = \beta$$

The solution set is then

$$\left(\alpha, \beta, \frac{4-3\alpha+\beta}{2}\right)$$

where $\alpha \in R$, $\beta \in R$.

It is noted that all these special cases arise when the determinant of the coefficients, $\Delta = 0$.

Exercise 3.5

Solve the following simultaneous equations

1. $\left.\begin{aligned} x-3y+z&=2\\ x+2y-z&=-1\\ 2x+y+2z&=6\end{aligned}\right\}$

2. $\left.\begin{aligned} x+y+z&=6\\ 2x&=y\\ x+y&=z\end{aligned}\right\}$

3. $\left.\begin{aligned} 2x-3y+z&=3\\ x+5y-z&=3\\ 7x+9y-z&=15\end{aligned}\right\}$

4. $\left.\begin{aligned} 2x-3y+z&=3\\ x+5y-z&=3\\ 6x-9y+3z&=9\end{aligned}\right\}$

5. $\left.\begin{aligned} x+7y+2z&=4\\ x-7y+2z&=2\\ 7x+z&=6\end{aligned}\right\}$

6. $\left.\begin{aligned} 7x-y-z&=64\\ 2x+5y-z&=59\\ x-3y-2z&=115\end{aligned}\right\}$

7. $\left.\begin{aligned} 9x-10y+z&=1\\ 3x+y-z&=-1\\ 7x-y+z&=1\end{aligned}\right\}$

8. $\left.\begin{aligned} 18x+3y+z&=49\\ x-7y+15z&=-54\\ 2x-4y+13z&=-25\end{aligned}\right\}$

9. $\left.\begin{aligned} 3x+2y-4z&=7\\ 2x-3y+5z&=2\\ 14x+31y-57z&=46\end{aligned}\right\}$

10. $\left.\begin{aligned} 18x-7y+5z&=4\\ 2x+7y-3z&=7\\ 16x-84y+40z&=62\end{aligned}\right\}$

11. Multiply together the two matrices

$$\begin{pmatrix} 1 & 3 & 2\\ 2 & -1 & 1\\ 3 & -2 & -1\end{pmatrix}\begin{pmatrix} 3 & -1 & 5\\ 5 & -7 & 3\\ -1 & 11 & -7\end{pmatrix}$$

Hence, or otherwise, solve the simultaneous equations

$$\begin{aligned} 3x-y+5z&=32\\ 5x-7y+3z&=-8\\ x-11y+7z&=20\end{aligned}$$

(*SMP Additional*)

Exercise 3.6 Miscellaneous

1. Use any suitable method to solve the equations.

(*a*) $$\left.\begin{aligned}10x - 4y + 8z &= -10\\ 3x + 2y + z &= 1\\ 5x - 2y + 4z &= 5\end{aligned}\right\}$$

(*b*) $$\left.\begin{aligned}15x + 9y - 3z &= 4\\ 10x + 6y - 2z &= -3\\ 35x + 21y - 7z &= 21\end{aligned}\right\}$$

(*c*) $$\left.\begin{aligned}x - 7y + 9z &= 26\\ 9x + 8y - z &= -1\\ 7x - 8y + 9z &= 33\end{aligned}\right\}$$

(*d*) $$\left.\begin{aligned}2x - y + 7z &= 4\\ 6x + 11y - 41z &= -4\\ 3x + 2y - 5z &= 2\end{aligned}\right\}$$

(*e*) $$\left.\begin{aligned}15x - 10y + 20z &= 30\\ x + y + z &= 0\\ 21x - 14y + 28z &= 42\end{aligned}\right\}$$

(*f*) $$\left.\begin{aligned}5x + 14y - 23z &= 2\\ x + 3y - 6z &= 9\\ x + 7y - 34z &= 180\end{aligned}\right\}$$

(*g*) $$\left.\begin{aligned}x + 2y - 3z - 4 &= 0\\ 2x - y + z - 5 &= 0\\ 3x + 2y + z - 14 &= 0\end{aligned}\right\}$$

(*h*) $$\begin{aligned}3x - 3y + 3z\\ = 6x + 3y + 8z\\ = -4\end{aligned}$$

2. Given the matrix

$$A = \begin{pmatrix} 1{\cdot}2 & 3 & 1\\ 0{\cdot}2 & 0 & -2\\ -0{\cdot}5 & 0{\cdot}1 & 2{\cdot}1\end{pmatrix}$$

and the matrix

$$B = \begin{pmatrix} 0{\cdot}2 & -6{\cdot}2 & -6\\ 0{\cdot}58 & 3{\cdot}02 & 2{\cdot}6\\ 0{\cdot}02 & -1{\cdot}62 & -0{\cdot}6\end{pmatrix}$$

Find the matrix AB. Hence, or otherwise, solve

$$\left.\begin{aligned}0{\cdot}2x - 6{\cdot}2y - 6z &= 12\\ 0{\cdot}58x + 3{\cdot}02y + 2{\cdot}6z &= 45\\ 0{\cdot}02x - 1{\cdot}62y - 0{\cdot}6z &= 78\end{aligned}\right\}$$

(*MEI Specimen*)

3. If two planes Π_1 and Π_2 have equations

$$\Pi_1 \equiv ax + by + cz = p$$
$$\Pi_2 \equiv dx + ey + fz = q$$

show that $\lambda\Pi_1 + \mu\Pi_2$ is a plane passing through the line of intersection of Π_1 and Π_2.

4. (i) $A = \begin{pmatrix} x & y\\ 0 & z\end{pmatrix}$. Find the most general form of the matrix B such that $AB = BA$ and prove that all such matrices B are of the form $\alpha I + \beta A$ where α and β are scalars.

(ii)

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & 1 \\ 1 & 4 & -1 \end{pmatrix} \text{ Calculate (i) det } A \quad \text{(ii) } A^{-1}.$$

Hence, or otherwise, solve the simultaneous equations:

$$\left.\begin{aligned} x + y + z &= 6 \\ 2x - y + z &= 3 \\ x + 4y - z &= 6 \end{aligned}\right\}$$

(*MEI Specimen*)

5. Find the inverse of

$$M = \begin{pmatrix} -1 & 3 & 4 \\ 5 & -2 & -3 \\ 2 & -3 & -4 \end{pmatrix}$$

Hence, or otherwise, solve the equations

$$M\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

where

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ -2 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}$$

(*MEI Specimen*)

6. Find the Inverse Matrix of

$$A = \begin{pmatrix} 2 & 3 & 1 \\ 1 & -1 & -2 \\ 3 & 2 & 3 \end{pmatrix}$$

Hence solve the simultaneous equations,

$$\left.\begin{aligned} 2x + 3y + z &= 7 \\ x - y - 2z &= 1 \\ 3x + 2y + 3z &= 4 \end{aligned}\right\}$$

(*MEI Specimen*)

7. Find the general solution of the system of equations

$$\begin{aligned} x + y - az &= b \\ 3x - 2y - z &= 1 \\ 4x - 3y - z &= 2 \end{aligned}$$

in each of the three cases (i) $a = 1$, $b = 9$; (ii) $a = 2$, $b = -3$; (iii) $a = 2$, $b = 0$. (*MEI Specimen*)

8. Find the value of a such that the three lines

$$\begin{aligned} 2x + 3y + 4 &= 0 \\ x - ay - 7 &= 0 \\ 3x + y - 1 &= 0 \end{aligned}$$

are concurrent. (*MEI Specimen*)

9. Find the inverse of the matrix

$$\begin{pmatrix} 1 & 3 & 2 \\ 5 & 2 & 7 \\ 1 & 2 & 1 \end{pmatrix}$$

What is the inverse of the matrix

$$\begin{pmatrix} 2 & 10 & 2 \\ 6 & 4 & 4 \\ 4 & 14 & 2 \end{pmatrix} ?$$

Solve the simultaneous equations

$$\begin{aligned} x + 3y + 2z &= 1 \\ 5x + 2y + 7z &= -2 \\ x + 2y + z &= 2 \end{aligned}$$

(*MEI Specimen*)

10. If

$$M = \begin{pmatrix} 2 & 3 & 6 \\ 6 & 2 & -3 \\ 3 & -6 & 2 \end{pmatrix},$$

form the product MM' and show that

$$M' = \frac{1}{49} M^{-1}$$

Without multiplying out, state the product $M'M$, giving reasons for your answer.

Hence, or otherwise, find the solution of the equations

$$\begin{aligned} 2x_1 + 3x_2 + 6x_3 &= 1 \\ 6x_1 + 2x_2 - 3x_3 &= 1 \\ 3x_1 - 6x_2 + 2x_3 &= 2 \end{aligned}$$

(*MEI Mathematics I*)

11. (*a*) Evaluate the product

$$\begin{pmatrix} 3 & -3 & 1 \\ -3 & 5 & -2 \\ 1 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{pmatrix}$$

Hence or otherwise solve the simultaneous equations

$$\begin{aligned} x + y + z &= 2 \\ x + 2y + 3z &= 6 \\ x + 3y + 6z &= 13 \end{aligned}$$

(*SMP Additional*)

HOMOGENEOUS EQUATIONS

Two equations in two unknowns

3.18 Consider the two equations

$$\left.\begin{aligned} ax + by &= 0 \\ cx + dy &= 0 \end{aligned}\right\}$$

If the corresponding lines are considered, they both pass through the origin, and hence the equations have the trivial solution

$$x = 0, \quad y = 0$$

Other solutions will also exist if the two equations represent the same line. In this case the gradients are the same.

$$\frac{a}{b} = \frac{c}{d}$$

$$\Rightarrow ad - bc = 0$$

$$\Rightarrow \begin{vmatrix} a & b \\ c & d \end{vmatrix} = 0$$

which is the condition for a special case as in **3.6.**

Three equations in three unknowns

3.19 Consider the equations

$$\begin{aligned} ax + by + cz &= 0 \\ dx + ey + fz &= 0 \\ gx + hy + iz &= 0 \end{aligned}$$

If the corresponding planes are considered, then clearly each plane passes through the origin.

The three equations therefore always have the trivial solution

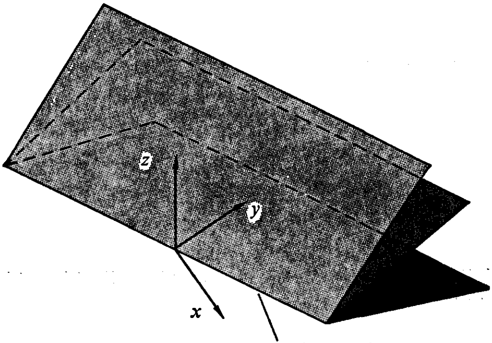
$$x = 0, \quad y = 0, \quad z = 0$$

Other solutions are only possible if one of three special cases apply, i.e. when

$$\Delta = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = 0 \text{ as in } \mathbf{3.14}$$

The three special cases

3.20 (*a*) The three planes may meet as the pages of a book in a line through the origin.

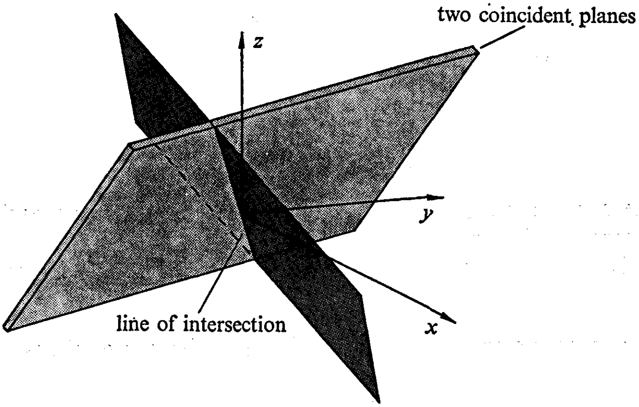


In this case, as with non-homogeneous equations,

$$\Delta = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = 0$$

The coordinates of any point on the line of intersection of the three planes will satisfy the three equations.

3.21 (*b*) Two planes may coincident and the third meet these in a line through the origin.

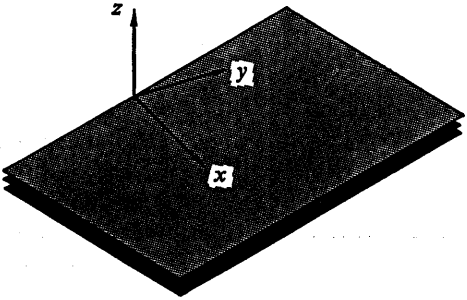


In this case again

$$\Delta = 0$$

The coordinates of any point on the line of intersection of the three planes will satisfy the three equations.

3.22 (*c*) The three planes may all be coincident, i.e. the three equations represent a single plane through the origin. Once again $\Delta = 0$. The coordinates of any point in this plane will satisfy the three equations.



3.23 *In general,* the equations

$$\begin{aligned} ax + by + cz &= 0 \\ dx + ey + fz &= 0 \\ gx + hy + iz &= 0 \end{aligned}$$

represent three planes through the origin. They always have the trivial solution $x = 0, y = 0, z = 0$.

Now non-trivial solutions can exist only if

$$\Delta = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = 0$$

and in this case there are three possible geometrical configurations.

Exercise 3.7

1. Determine the values of λ for which the equations

$$\left.\begin{aligned} (\lambda - 2)x - 2y - z &= 0 \\ x - (\lambda - 3)y + z &= 0 \\ -x - 2y + (\lambda - 2)z &= 0 \end{aligned}\right\}$$

have a solution other than the trivial one.

2. Express one of the equations

$$\begin{aligned} 2x + y + z &= 0 \\ 2x + 2y + 3z &= 0 \\ 6x + 5y + 7z &= 0 \end{aligned}$$

as a linear combination of the other two, and find the general solution of the equations.

Also show that if any one of the coefficients is altered, the resulting equations will have a non-trivial solution. *(University of Leicester)*

3. Prove that the values of k for which the three simultaneous equations

$$\begin{aligned} x_1 \qquad\quad + x_3 &= kx_1 \\ -2x_1 + 2x_2 \qquad\quad &= kx_2 \\ -\ x_1 + \ x_2 - x_3 &= kx_3 \end{aligned}$$

have solutions *other* than $x_1 = x_2 = x_3 = 0$ satisfy the cubic equation $k^3 - 2k^2 + 2 = 0$.

Prove that one root of this cubic equation lies between -1 and 0 and find this root correct to two decimal places by using an appropriate iterative method. *(MEI Mathematics I)*

Choleski's method

3.24 This is one of the quickest methods of solving three simultaneous equations in three unknowns. The method is demonstrated with a worked example.

Consider the equations written in matrix form as

$$\begin{pmatrix} -3 & 2 & 1 \\ 9 & -5 & -2 \\ -5 & 3 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 8 \\ -21 \\ 17 \end{pmatrix} \qquad \text{(i)}$$

$$\text{i.e.} \qquad A \qquad \mathbf{x} = \mathbf{b}$$

The solution of such equations by the previously explained matrix method* involves finding A^{-1}, the inverse of A. Choleski's method does not use an inverse. Instead use is made of the fact that an equation such as

$$\begin{pmatrix} -3 & 0 & 0 \\ 9 & 1 & 0 \\ -5 & -\frac{1}{3} & -\frac{1}{3} \end{pmatrix}\begin{pmatrix} p \\ q \\ r \end{pmatrix} = \begin{pmatrix} 8 \\ -21 \\ 17 \end{pmatrix} \qquad \text{(ii)}$$

where the matrix of the coefficients has zeros above the leading diagonal, is easy to solve. In this case multiplying out gives

$$\begin{aligned} -3p &= \quad 8 \Rightarrow p = -\tfrac{8}{3} \\ 9p + q &= -21 \Rightarrow q = \quad 3 \\ -5p - \tfrac{1}{3}q - \tfrac{1}{3}r &= \quad 17 \Rightarrow r = -14 \end{aligned}$$

---

* See 3.8.

3.25 The method involves looking for two matrices L and U so that $LU = A$, and where

L is a matrix with zeros above the leading diagonal, as above, and
U is a matrix with zeros below the leading diagonal, which is arranged to be $\{1, 1, 1\}$.

Now to find L and U, it is noted that

$$\begin{pmatrix} -3 & 0 & 0 \\ 9 & l & 0 \\ -5 & m & n \end{pmatrix}\begin{pmatrix} 1 & u & v \\ 0 & 1 & w \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -3 & 2 & 1 \\ 9 & -5 & -2 \\ -5 & 3 & 1 \end{pmatrix}$$

Multiplying out

$$\begin{aligned} -3u &= 2 \Rightarrow u = -\tfrac{2}{3} \\ -3v &= 1 \Rightarrow v = -\tfrac{1}{3} \\ 9u + l &= -5 \Rightarrow l = 1 \\ 9v + lw &= -2 \Rightarrow w = 1 \\ -5u + m &= 3 \Rightarrow m = -\tfrac{1}{3} \\ -5v + mw + n &= 1 \Rightarrow n = -\tfrac{1}{3} \end{aligned}$$

Hence

$$\begin{array}{ccc} L & U & = \quad A \end{array}$$

$$\begin{pmatrix} -3 & 0 & 0 \\ 9 & 1 & 0 \\ -5 & -\tfrac{1}{3} & -\tfrac{1}{3} \end{pmatrix}\begin{pmatrix} 1 & -\tfrac{2}{3} & -\tfrac{1}{3} \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -3 & 2 & 1 \\ 9 & -5 & -2 \\ -5 & 3 & 1 \end{pmatrix}$$

Now the original matrix equation was

$$A\mathbf{x} = \mathbf{b}$$
$$\Rightarrow LU\mathbf{x} = \mathbf{b}$$

But if $U\mathbf{x} = \mathbf{p}$, then

$$L\mathbf{p} = \mathbf{b}$$

But from (ii)

$$\mathbf{p} = \begin{pmatrix} p \\ q \\ r \end{pmatrix} = \begin{pmatrix} -\tfrac{8}{3} \\ 3 \\ -14 \end{pmatrix}$$

Hence $U\mathbf{x} = \mathbf{p}$

$$\Rightarrow \begin{pmatrix} 1 & -\tfrac{2}{3} & -\tfrac{1}{3} \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -\tfrac{8}{3} \\ 3 \\ -14 \end{pmatrix}$$

Multiplying out

$$\begin{aligned} z &= -14 \\ y + z &= 3 \Rightarrow y = 17 \\ x - \tfrac{2}{3}y - \tfrac{1}{3}z &= -\tfrac{8}{3} \Rightarrow x = 4 \end{aligned}$$

Hence the solution set is

$$\mathbf{x} = \begin{pmatrix} 4 \\ 17 \\ -14 \end{pmatrix}$$

Exercise 3.8

Use Choleski's Method to solve the following equations:

1. $\left.\begin{aligned} x - 2y + z &= -1 \\ y - z &= 1 \\ 3x + y - 2z &= 4 \end{aligned}\right\}$ 2. $\left.\begin{aligned} 2x + y - z &= 0 \\ x - y + z &= 6 \\ x + 2y + z &= 3 \end{aligned}\right\}$

Some special cases

3.26 The special cases are demonstrated by the following examples:

Example 1

$$\left.\begin{aligned} x + y + z &= 3 \\ 2x - y + 3z &= 4 \\ 4x + y + 5z &= 10 \end{aligned}\right\}$$

$$\begin{array}{ccccc} A & & \mathbf{x} & = & \mathbf{b} \\ \begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & 3 \\ 4 & 1 & 5 \end{pmatrix} & & \begin{pmatrix} x \\ y \\ z \end{pmatrix} & = & \begin{pmatrix} 3 \\ 4 \\ 10 \end{pmatrix} \end{array}$$

To find L and U

$$\begin{array}{ccccc} L & & U & = & A \\ \begin{pmatrix} 1 & 0 & 0 \\ 2 & l & 0 \\ 4 & m & n \end{pmatrix} & & \begin{pmatrix} 1 & u & v \\ 0 & 1 & w \\ 0 & 0 & 1 \end{pmatrix} & = & \begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & 3 \\ 4 & 1 & 5 \end{pmatrix} \end{array}$$

Hence

$$u = 1$$

$$v = 1$$

$$\begin{aligned} 2u + l &= -1 \Rightarrow l = -3 \\ 2v + lw &= 3 \Rightarrow w = -\tfrac{1}{3} \\ 4u + m &= 1 \Rightarrow m = -3 \\ 4v + wm + n &= 5 \Rightarrow n = 0 \end{aligned}$$

Hence

$$\begin{array}{ccc} L & U & = & A \\ \begin{pmatrix} 1 & 0 & 0 \\ 2 & -3 & 0 \\ 4 & -3 & 0 \end{pmatrix} & \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -\frac{1}{3} \\ 0 & 0 & 1 \end{pmatrix} & = & \begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & 3 \\ 4 & 1 & 5 \end{pmatrix} \end{array}$$

Now if

$$\begin{array}{cccc} L & \mathbf{p} & = & \mathbf{b} \\ \begin{pmatrix} 1 & 0 & 0 \\ 2 & -3 & 0 \\ 4 & -3 & 0 \end{pmatrix} & \begin{pmatrix} p \\ q \\ r \end{pmatrix} & = & \begin{pmatrix} 3 \\ 4 \\ 10 \end{pmatrix} \end{array}$$

then

$$p = 3$$

$$2p - 3q = 4 \Rightarrow q = \tfrac{2}{3}$$

$$4p - 3q = 10$$

which is satisfied by the above values of p and q. r is undetermined.

Hence

$$\mathbf{p} = \begin{pmatrix} 3 \\ \frac{2}{3} \\ r \end{pmatrix}$$

Hence

$$U\mathbf{x} = \mathbf{p}$$

$$\Rightarrow \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -\frac{1}{3} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3 \\ \frac{2}{3} \\ r \end{pmatrix}$$

$$\Rightarrow z = r$$

$$y - \tfrac{1}{3}z = \tfrac{2}{3} \Rightarrow y = \frac{2+r}{3}$$

$$x + y + z = 3 \Rightarrow x = 3 - \frac{2+r}{3} - r = \frac{7-4r}{3}$$

Hence the solution set is the infinite set of points

$$\left(\frac{7-4r}{3}, \frac{2+r}{3}, r\right)$$

for all real values of r.*

---

* cf. p. 51.

Example 2

$$\left.\begin{aligned} 2x - 3y + z &= 4 \\ 4x - 6y + 2z &= 5 \\ x - y - z &= 1 \end{aligned}\right\}$$

$$\overset{A}{\begin{pmatrix} 2 & -3 & 1 \\ 4 & -6 & 2 \\ 1 & -1 & -1 \end{pmatrix}} \overset{\mathbf{x}}{\begin{pmatrix} x \\ y \\ z \end{pmatrix}} \overset{=}{=} \overset{\mathbf{b}}{\begin{pmatrix} 4 \\ 5 \\ 1 \end{pmatrix}}$$

$$\overset{L}{\begin{pmatrix} 2 & 0 & 0 \\ 4 & l & 0 \\ 1 & m & n \end{pmatrix}} \overset{U}{\begin{pmatrix} 1 & u & v \\ 0 & 1 & w \\ 0 & 0 & 1 \end{pmatrix}} = \overset{A}{\begin{pmatrix} 2 & -3 & 1 \\ 4 & -6 & 2 \\ 1 & -1 & -1 \end{pmatrix}}$$

Hence

$$\begin{aligned} 2u &= -3 \Rightarrow u = -\tfrac{3}{2} \\ 2v &= 1 \Rightarrow v = \tfrac{1}{2} \\ 4u + l &= -6 \Rightarrow l = 0 \end{aligned}$$

$4v + lw = 2$ which is satisfied by any value of w

$$\begin{aligned} u + m &= -1 \Rightarrow m = \tfrac{1}{2} \\ v + wm + n &= -1 \Rightarrow n = -\tfrac{1}{2} - \tfrac{1}{2}w \end{aligned}$$

$$\overset{L}{\begin{pmatrix} 2 & 0 & 0 \\ 4 & 0 & 0 \\ 1 & \frac{1}{2} & -\frac{1}{2} - \frac{1}{2}w \end{pmatrix}} \overset{U}{\begin{pmatrix} 1 & -\frac{3}{2} & \frac{1}{2} \\ 0 & 1 & w \\ 0 & 0 & 1 \end{pmatrix}} \overset{=}{=} \overset{A}{\begin{pmatrix} 2 & -3 & 1 \\ 4 & -6 & 2 \\ 1 & -1 & -1 \end{pmatrix}}$$

Now if $L\mathbf{p} = \mathbf{b}$

$$\begin{pmatrix} 2 & 0 & 0 \\ 4 & 0 & 0 \\ 1 & \frac{1}{2} & -\frac{1}{2} - \frac{1}{2}w \end{pmatrix} \begin{pmatrix} p \\ q \\ r \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \\ 1 \end{pmatrix}$$

Hence

$$\begin{aligned} 2p &= 4 \Rightarrow p = 2 \\ 4p &= 5 \Rightarrow p = \tfrac{5}{4} \end{aligned}$$

which is inconsistent. Hence there is no solution.

Exercise 3.9

Use Choleski's method to find the solution, if it exists, of the following equations.

1. $$\left.\begin{aligned} x + 3y + 2z &= 1 \\ 5x + 2y + 7z &= -2 \\ x + 2y + z &= 2 \end{aligned}\right\}$$

2. $$\left.\begin{aligned} 2x + y - 3z &= 4 \\ 3x - 4y + 5z &= 2 \\ 5x - 3y + 2z &= 6 \end{aligned}\right\}$$

3. $$\left.\begin{aligned} 6x - 2y + 4z &= 8 \\ -9x + 3y - 6z &= -12 \\ 36x - 12y + 24z &= 48 \end{aligned}\right\}$$

4. $$\left.\begin{aligned} 4x - 6y + 2z &= 5 \\ -6x + 9y - 3z &= 15 \\ x + y + z &= 2 \end{aligned}\right\}$$

Classification of the solutions of linear equations*

3.11 *Definitions.* The rank of a matrix is the order† of its largest square 'sub'-matrix whose determinant is non-zero. The rank of a null matrix is zero. A set of linear equations with one or more solutions is called a consistent set. If there is no solution the equations are inconsistent. A set of linear equations $A\mathbf{x} = \mathbf{b}$ is consistent if and only if the rank of the augmented matrix (containing A with the additional column $\mathbf{b}$ and written $A/\mathbf{b}$ is equal to the rank of A. This common value of the rank is called the rank of the set r.

A consistent set of linear equations $A\mathbf{x} = \mathbf{b}$ is n unknowns has a unique solution if and only if $r = n$. Given a set of linear equations $A\mathbf{x} = \mathbf{b}$ in n unknowns, then

either (1) they are inconsistent and rank $A <$ rank $A/\mathbf{b}$

or (2) they are consistent and rank $A =$ rank $A/\mathbf{b} = r$

In this case we have

(*a*) a unique solution if $r = n$

or (*b*) an infinite number of solutions if $r < n$.

Example 1

Consider the set of linear equations

$$\begin{pmatrix} 1 & 3 & 2 \\ 5 & 2 & 7 \\ 1 & 2 & 1 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}$$

i.e. $$A = \begin{pmatrix} 1 & 3 & 2 \\ 5 & 2 & 7 \\ 1 & 2 & 1 \end{pmatrix} \quad \text{and} \quad A/\mathbf{b} = \begin{pmatrix} 1 & 3 & 2 & 1 \\ 5 & 2 & 7 & -2 \\ 1 & 2 & 1 & 2 \end{pmatrix}$$

Now the rank of A must be 3 or less, and we only need to find one non-zero determinant of order 3 to show that the rank of A is 3. Now $|A| = 10$. Hence the rank of $A =$ rank of $A/\mathbf{b} = 3$, and the set has a unique solution.

---

[*] This is a brief note only and is a guide to further work; it is in no way a justification.

[†] The order of a square matrix or determinant is not usually given as 2 × 2 or 3 × 3 but left as 2 or 3.

Example 2

Consider the set of linear equations

$$\begin{pmatrix} 2 & 1 & -3 \\ 3 & -4 & 5 \\ 5 & -3 & 2 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 4 \\ 2 \\ 6 \end{pmatrix}$$

i.e. $A = \begin{pmatrix} 2 & 1 & -3 \\ 3 & -4 & 5 \\ 5 & -3 & 2 \end{pmatrix}$ and $A/\mathbf{b} = \begin{pmatrix} 2 & 1 & -3 & 4 \\ 3 & -4 & 5 & 2 \\ 5 & -3 & 2 & 6 \end{pmatrix}$

Now $|A| = 0$, hence the rank of A is less than 3.

Consider $\begin{vmatrix} 2 & 1 \\ 3 & -4 \end{vmatrix} = -11$. Hence the rank of $A = 2$.

Consider the 3×3 submatrices of $A/\mathbf{b}$, i.e.

$$\begin{pmatrix} 2 & 1 & -3 \\ 3 & -4 & 5 \\ 5 & -3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & -3 & 4 \\ -4 & 5 & 2 \\ -3 & 2 & 6 \end{pmatrix}, \begin{pmatrix} 2 & -3 & 4 \\ 3 & 5 & 2 \\ 5 & 2 & 6 \end{pmatrix}, \begin{pmatrix} 2 & 1 & 4 \\ 3 & -4 & 2 \\ 5 & -3 & 6 \end{pmatrix}$$

Each of these have determinant zero.

Hence rank A = rank $A/\mathbf{b} = 2$.
Since $r = 2$ and $n = 3$ we have an infinite number of solutions.

Example 3

$$\begin{aligned} x - 2y + 2z &= 3 \\ 2x - 4y + 4z &= 1 \\ 3x - 3y - 3z &= 4 \end{aligned}$$

i.e. $$\begin{pmatrix} 1 & -2 & 2 \\ 1 & -4 & 4 \\ 3 & -3 & -3 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix}$$

$|A| = 0$, but $\begin{vmatrix} 1 & -2 \\ 1 & -4 \end{vmatrix} = -2$ and so the rank of A is 2.

Consider $$A/\mathbf{b} = \begin{pmatrix} 1 & -2 & 2 & 3 \\ 1 & -4 & 4 & 1 \\ 3 & -3 & -3 & 4 \end{pmatrix}$$

Now $\begin{vmatrix} -2 & 2 & 3 \\ -4 & 4 & 1 \\ -3 & -3 & 4 \end{vmatrix} = 60$. Hence the rank of $A/\mathbf{b}$ is 3.

Since rank $A <$ rank $A/\mathbf{b}$ there is no solution.

Exercise 3.10 Miscellaneous

1. Find the vertices of the tetrahedron given by the intersection of the planes

$$\begin{aligned} x + 2y + 3z &= 0 \\ 2x + y - z &= 0 \\ x + 3y - z &= 0 \\ x + y + z &= 1 \end{aligned}$$

2. Obtain general solutions for the equations

$$\begin{aligned} x + \alpha y + \beta z + \gamma &= 0 \\ x + \beta y + \gamma z + \alpha &= 0 \\ x + \gamma y + \alpha z + \beta &= 0 \end{aligned}$$

in the cases

(i) α, β, γ all different.
(ii) $\alpha \neq \beta = \gamma$.
(iii) $\alpha = \beta = \gamma$. *(CS)*

3. Solve the following equations completely

$$\begin{aligned} 2x + 5y + z &= a \\ x + 3y + az &= 1 \\ 3x + 8y + bz &= c \end{aligned}$$

In particular, for what values of a, b, c have these equations

(i) no solutions,
(ii) more than one solution? *(CS)*

4. For what values of a, b and c are the following equations consistent?

$$\begin{aligned} x + y + z &= 1 \\ ax + by + cz &= 0 \\ a^2x + b^2y + c^2z &= 0 \end{aligned}$$

Solve them completely when they are consistent. *(CS)*

5. Investigate the solutions of the equations

$$3ax - 6y + 18z = b, \quad 4x + 5y - z = -2, \quad x - 7y - z = 3$$

in the cases

(i) $a = 3$, $b = 21$.
(ii) $a = -17$, $b = 1$. *(MT)*

6. Investigate the solutions of the equations

$$a^2x + ay + z = a^2, \quad ax + y + bz = 1, \quad a^2bx + y + bz = b$$

for all possible values of a, b.

(MT)

7. If k is given to be real, show that the equations

$$\begin{aligned} kx + y + z &= 2 \\ 2x + ky + z &= 1 \\ -4x + y + kz &= 1 \end{aligned}$$

have a unique solution, in which $x = 2k/(k^2 + k + 2)$ unless k takes one exceptional value. Also show that in the exceptional case there is no solution. *(University of Leicester)*

8. Find when the equations

$$\begin{aligned} x + y - 2z &= 0 \\ ax + by + cz &= 0 \\ bx + cy + az &= d \end{aligned}$$

are consistent, and solve them completely when they are consistent.

(University of Leicester)

9. The numbers a, b, c, α, β, γ are all real. Find (do not merely quote) the condition on a, b, c, that the equations

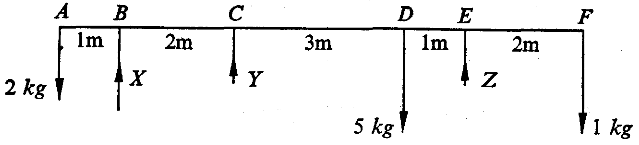
$$\begin{aligned} ax + by + cz &= \alpha \\ bx + cy + az &= \beta \\ cx + ay + bz &= \gamma \end{aligned}$$

should not have a unique solution for x, y, z.

Separate the condition into two cases, and in each case find the conditions on α, β, γ so there should be infinitely many solutions.

(O. & C.)

10.



A light beam AF is loaded with weights at A, D and F. The beam is supported by posts at B, C and E. By resolving and taking moments form three equations, and solve these to find the thrusts X, Y and Z in the posts necessary to maintain equilibrium.

Determinants and Curve-fitting

3.12 Determinants can be used to find the equation of a curve or surface passing through a given set of points. To illustrate the method in its simplest form, consider the equation of a straight line.

Example 1

A straight line has the form

$$Ax + By + C = 0$$

and uses three constants. Its equation is equivalent to

$$Px + Qy + 1 = 0$$

and this form requires only two constants. Hence any straight line can be put in a form using two constants only. A straight line then has two *degrees of freedom*, and it is noted that two distinct points determine a line.

To find the equation of the line passing through a given pair of points it is necessary to find the constants P and Q. If the line is to pass through the fixed points (x_1, y_1) and (x_2, y_2) and also the general point (x, y) then

$$\begin{aligned} Px + Qy + 1 &= 0 \\ Px_1 + Qy_1 + 1 &= 0 \\ Px_2 + Qy_2 + 1 &= 0 \end{aligned}$$

From these three equations P and Q can be eliminated.

Consider the equations as a set of three homogeneous equations in three unknowns, P, Q, R.

$$\begin{aligned} Px + Qy + R &= 0 \\ Px_1 + Qy_1 + R &= 0 \\ Px_2 + Qy_2 + R &= 0 \end{aligned}$$

where R = 1.

The equations cannot have the trivial solution P = 0, Q = 0, R = 0 since R = 1. Hence the equations can have only a non-trivial solution. This solution exists only if

$$\begin{vmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} = 0$$

This equation is now the condition that the line $Px + Qy + 1 = 0$ passes through the variable point (x, y) and the given fixed points (x_1, y_1) and (x_2, y_2). Hence the equation gives, in determinant form, the equation of the straight line required.

Worked Example

The equation of the straight line through (1, 2) and (−7, 3) is

$$\begin{vmatrix} x & y & 1 \\ 1 & 2 & 1 \\ -7 & 3 & 1 \end{vmatrix} = 0$$

$$\Rightarrow x + 8y - 17 = 0$$

The same technique can be applied to finding the equation of any algebraic curve or surface through a set of fixed points.

Example 2

A general circle has an equation of the form

$$x^2 + y^2 + 2gx + 2fy + c = 0$$

This equation is equivalent to

$$\mathrm{P}(x^2 + y^2) + \mathrm{Q}x + \mathrm{R}y + 1 = 0, \text{ provided } \mathrm{P} \neq 0$$

The equation now involves only three constants. This means that three distinct points are necessary to define a circle. Geometrically this can be justified since a unique circumcircle can be constructed around the triangle formed by the three given points. If the circle is to pass through the fixed points (x_1, y_1), (x_2, y_2), and (x_3, y_3) and the variable point (x, y), then it is necessary to solve the following equations for P, Q and R

$$\begin{aligned} \mathrm{P}(x^2 + y^2) + \mathrm{Q}x + \mathrm{R}y + 1 &= 0 \\ \mathrm{P}(x_1^2 + y_1^2) + \mathrm{Q}x_1 + \mathrm{R}y_1 + 1 &= 0 \\ \mathrm{P}(x_2^2 + y_2^2) + \mathrm{Q}x_2 + \mathrm{R}y_2 + 1 &= 0 \\ \mathrm{P}(x_3^2 + y_3^2) + \mathrm{Q}x_3 + \mathrm{R}y_3 + 1 &= 0 \end{aligned}$$

Again these can be considered as a set of four homogeneous equations in four unknowns, P, Q, R, S where S = 1. Again since S = 1, a non-trivial solution is required.

A non-trivial solution exists only if

$$\begin{vmatrix} x^2 + y^2 & x & y & 1 \\ x_1^2 + y_1^2 & x_1 & y_1 & 1 \\ x_2^2 + y_2^2 & x_2 & y_2 & 1 \\ x_3^2 + y_3^2 & x_3 & y_3 & 1 \end{vmatrix} = 0$$

This equation gives the condition that the four points (x_1, y_1), (x_2, y_2), (x_3, y_3) and (x, y) lie on a circle. Hence it gives the equation of the circle required. It was required, however, that $P \neq 0$. Algebraically this means that the cofactor of $x^2 + y^2$ cannot equal zero

i.e.
$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} \neq 0$$

Geometrically it is seen from the first example that this is the condition that (x_1, y_1), (x_2, y_2) and (x_3, y_3) do not lie on a straight line. The same technique can now be applied to other problems.

Worked Example

Find the equation of a rectangular hyperbola with its asymptotes parallel to the coordinate axes, and that passes through the points $(6, 6)$, $(5, 10)$, $(-3, -6)$.

The equation of a rectangular hyperbola with centre the origin and with its asymptotes along the coordinate axes is

$$xy = c$$

If the centre is not at the origin but is at (h, k) then the rectangular hyperbola would have an equation

$$(x - h)(y - k) = c$$
$$\Rightarrow xy - kx - hy + hk - c = 0$$

This equation has the form

$$Pxy + Qx + Ry + 1 = 0$$

where $P \neq 0$

Hence the equation of the hyperbola required is

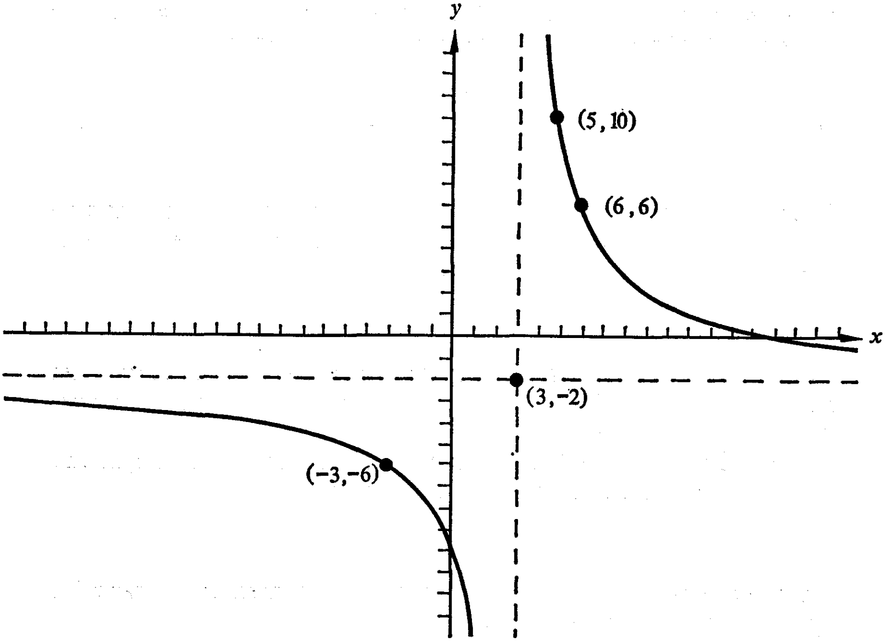
$$\begin{vmatrix} xy & x & y & 1 \\ 36 & 6 & 6 & 1 \\ 50 & 5 & 10 & 1 \\ 18 & -3 & -6 & 1 \end{vmatrix} = 0$$

$$\Rightarrow \begin{vmatrix} 6 & 6 & 1 \\ 5 & 10 & 1 \\ -3 & -6 & 1 \end{vmatrix} xy - \begin{vmatrix} 36 & 6 & 1 \\ 50 & 10 & 1 \\ 18 & -6 & 1 \end{vmatrix} x + \begin{vmatrix} 36 & 6 & 1 \\ 50 & 5 & 1 \\ 18 & -3 & 1 \end{vmatrix} y - \begin{vmatrix} 36 & 6 & 6 \\ 50 & 5 & 10 \\ 18 & -3 & -6 \end{vmatrix} = 0$$

$$\Rightarrow 48xy + 96x - 144y - 1440 = 0$$

$$\Rightarrow xy + 2x - 3y - 30 = 0$$

$$\Rightarrow (x - 3)(y + 2) = 24$$



Note that the coefficient of $xy \neq 0$

i.e. $$\begin{vmatrix} 6 & 6 & 1 \\ 5 & 10 & 1 \\ -3 & -6 & 1 \end{vmatrix} \neq 0$$

This is the condition that the points (6, 6), (5, 0), (−3, −6) do not lie on a straight line.

Worked Example

Find the equation of an ellipsoid of revolution about the z-axis, with centre the origin, and with its axes parallel to the coordinate axes, such that it passes through (0, 4, 6) and (3, 4, 0).

The form of the equation required is

$$\frac{x^2}{a^2} + \frac{y^2}{a^2} + \frac{z^2}{b^2} = 1$$

$$\Rightarrow b^2(x^2 + y^2) + a^2z^2 - a^2b^2 = 0$$

$$\Rightarrow \mathrm{P}(x^2 + y^2) + \mathrm{Q}z^2 + 1 = 0$$

Hence the equation required is

$$\begin{vmatrix} x^2+y^2 & z^2 & 1 \\ 0^2+4^2 & 6^2 & 1 \\ 3^2+4^2 & 0^2 & 1 \end{vmatrix} = 0.$$

$$\Rightarrow \begin{vmatrix} x^2+y^2 & z^2 & 1 \\ 16 & 36 & 1 \\ 25 & 0 & 1 \end{vmatrix} = 0$$

$$\Rightarrow 36(x^2+y^2)+9z^2-900=0$$

$$\Rightarrow \frac{x^2}{5^2}+\frac{y^2}{5^2}+\frac{z^2}{10^2}=1$$

Exercise 3.11

1. Use the determinant method to find the equation of a line passing through

(*a*) $(2, -1)$ and $(-2, 3)$
(*b*) $(1, 1)$ and $(2, 1)$
(*c*) $(a, 0)$ and $(0, b)$

2. Use the determinant method to find, where possible, the equation of a circle passing through

(*a*) $(9, -7), (6, 2), (-3, 1)$
(*b*) $(0, 8), (6, 0), (6, 8)$
(*c*) $(1, 1), (4, 7), (-3, -7)$
(*d*) $(0, 0), (-1, 1), (6, 8)$

3. Use determinants to find, where possible, the equation of a plane passing through

(*a*) $(2, 0, 0), (0, 3, 0), (0, 0, 4)$
(*b*) $(4, 1, 8), (3, -1, 6), (1, -5, 2)$
(*c*) $(14, 4, 11), (1, 2, 41), (33, 1, 7)$

4. A parabola P_1, with vertex the origin, and with its axis lying along the x-axis has an equation of the form

$$y^2 = 4ax$$

What is the form of the equation of a parabola P_2 whose axis is parallel to that of P_1, but whose vertex is the point (h, k)? Express in determinant form the equation of a parabola with its axis parallel to the x-axis, and that passes through the three points

$$(x_1, y_1), \quad (x_2, y_2), \quad (x_3, y_3)$$

Find the equation of such a parabola passing through

$$(-1, 3), \quad (-4, -9), \quad (-1, -5)$$

5. Find in determinant form the equation of a parabola with its axis parallel to the y-axis and passing through three fixed points,

$$(x_1, y_1), \quad (x_2, y_2) \quad \text{and} \quad (x_3, y_3)$$

Find the equation of such a parabola passing through

$$(-1, 4), \quad (0, 5), \quad (-4, 13)$$

6. Find in determinant form the equation of a sphere, with its centre not at the origin, passing through each of a set of four fixed points of the form

$$(x_r, y_r, z_r), \quad r = 1, 2, 3, 4$$

What is the geometrical significance of the non-vanishing co-factor of the term in $(x^2 + y^2 + z^2)$?

7. Find the equation of the hyperboloid of one sheet of the form

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$$

that passes through the three given points

$$(3, 4, 12), \quad (0, 4, 0), \quad (5, 0, 16)$$

8. Find the equation of the central conic, centre the origin, with axes along the coordinate axes, and that passes through

(*a*) (26, 5), (−40, 16)
(*b*) (7, 48), (20, 30)
(*c*) (15, 12), (41, 40)
(*d*) (4, −2), (−6, −3)
(*e*) (15,20), (7, 24)

9. (*a*) The general second-degree equation in two variables is a conic, which takes the form

$$ax^2 + 2hxy + by^2 + 2gx + 2fy + c = 0$$

How many points are necessary to determine a conic? If these fixed points are of the form (x_r, y_r) where $r = 1, 2, \ldots$, write in determinant form the equation of a conic passing through these fixed points.

(*b*) Find the equation of a conic passing through

$$(0, 3);\quad (0, -3);\quad (-3, 0);\quad (1, -4);\quad (1, 3)$$

10. How many fixed points determine a general third-degree equation in two variables? Find in determinant form the equation of a cubic curve passing through these points.

11. How many fixed points determine a general second-degree equation in three variables. Find in determinant form the equation of a quadric surface passing through these points.

12. Write in determinant form the equation of a four dimensional linear surface passing through a set of fixed points. What condition must be satisfied for the surface to be determined?

4
Matrix Transformations

4.1 The ideas of simple matrix transformations will all have been met with before. (For many examples and details see *Matrices I*, *Matrices II* and *Exercises II* in the Contemporary School Mathematics series.) We only give here a brief reminder of these ideas.

Transformations of the plane using 2×2 matrices

4.2 If a column vector is pre-multiplied by a 2 × 2 matrix then the vector will be 'transformed' into a new one. If

$$\begin{pmatrix} x \\ y \end{pmatrix} \quad \text{is transformed to} \quad \begin{pmatrix} x' \\ y' \end{pmatrix}$$

we also speak of the 'transformation' or the mapping of the corresponding point, i.e. (x, y) to (x', y').

4.3 We consider the transformation of the points (1, 0) and (0, 1), i.e. the transformation of the vectors

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Now, any point (a, b) has associated with it the vector $\begin{pmatrix} a \\ b \end{pmatrix}$, which may be expressed as

$$a\begin{pmatrix} 1 \\ 0 \end{pmatrix} + b\begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Now if M is a 2 × 2 matrix, then we can transform every point of the plane using this matrix. Now*

$$\begin{aligned} M\begin{pmatrix} a \\ b \end{pmatrix} &= M\left[a\begin{pmatrix} 1 \\ 0 \end{pmatrix} + b\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right] \\ &= aM\begin{pmatrix} 1 \\ 0 \end{pmatrix} + bM\begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{aligned}$$

---

* Using ideas of matrix addition and multiplication as in Chapter 1.

Now

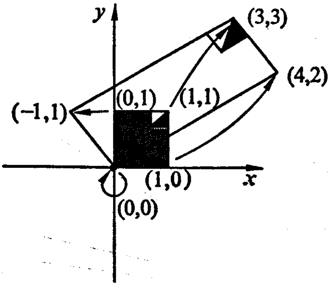
$$M\begin{pmatrix}1\\0\end{pmatrix} \quad \text{and} \quad M\begin{pmatrix}0\\1\end{pmatrix}$$

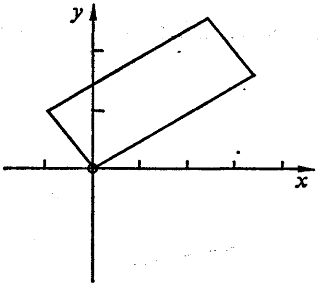
are column vectors representing the images of the points (1, 0) and (0, 1) under the transformation. Hence if we can find the images of the points (1, 0) and (0, 1) we can find the image of any other point (a, b). To do this it is easiest to use a unit square.*

4.4 *Example*. Investigate the transformation of the (x, y) plane represented by the matrix

$$\begin{pmatrix}4 & -1\\2 & 1\end{pmatrix}$$

We consider, first of all, the transformation of the points lying at the vertices of the *unit square*. This will indicate the transformation of the whole plane.





$$\begin{pmatrix}4 & -1\\2 & 1\end{pmatrix}\begin{pmatrix}0\\0\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$$ i.e. (0, 0) is transformed into (0, 0)

$$\begin{pmatrix}4 & -1\\2 & 1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}4\\2\end{pmatrix}$$ i.e. (1, 0) is transformed into (4, 2)

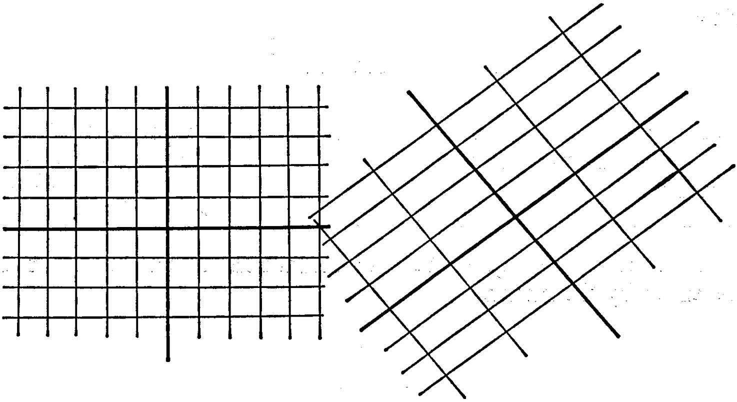
$$\begin{pmatrix}4 & -1\\2 & 1\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}-1\\1\end{pmatrix}$$ i.e. (0, 1) is transformed into (−1, 1)

$$\begin{pmatrix}4 & -1\\2 & 1\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}3\\3\end{pmatrix}$$ i.e. (1, 1) is transformed into (3, 3)

The unit square has become distorted as shown. Hence we expect the transformation of the plane to be something like this!

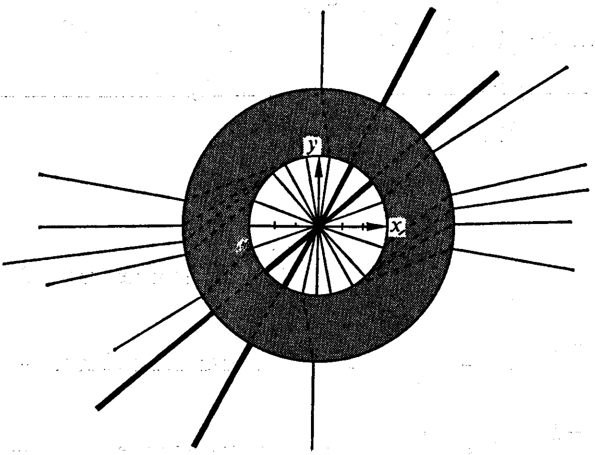
---

* The unit square has vertices (0, 0), (1, 0), (0, 1) and (1, 1).



The original plane The transformed plane

4.5 We can, rather than considering the transformation of the vertices of the unit square, consider the transformation of lines (through the origin), i.e. considering $\begin{pmatrix} x \\ y \end{pmatrix}$ as a position vector.



Consider points of the line $y = x$, i.e. (1, 1), (2, 2), (3, 3). These are transformed into (3, 3), (6, 6), (9, 9)—they still lie on the same line. In the same way we can consider the lines $y = -x$, $y = 2x$, $y = -2x$, $y = \frac{1}{2}x$, $y = -\frac{1}{2}x$, $y = 4x$, $y = -4x$.

The inner area represents the original whilst the area outside the ring shows the plane onto which the position vectors are transformed.

IMPORTANT: Note that the points of the lines which do not remain the same under the transformation 'close up' on one of those that does;

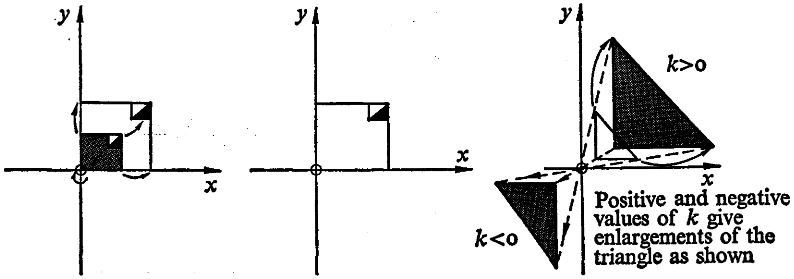
in fact, the whole plane 'closes up' in the direction of this line or lines. The lines which remain the same under the transformation are in the directions of the **eigenvectors** of the transformation matrix. These eigenvectors are such that

$$\begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = k\begin{pmatrix} x \\ y \end{pmatrix}, \qquad \text{where } k \in R$$

$\Bigg[$Here $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ are eigenvectors—corresponding to the fact that $y = x$ and $y = 2x$ remain the same lines under the transformation.$\Bigg]$ Eigenvectors are dealt with in detail in **4.22.**

Examples of other transformations

4.6 (*a*) *Enlargements centred on the origin*



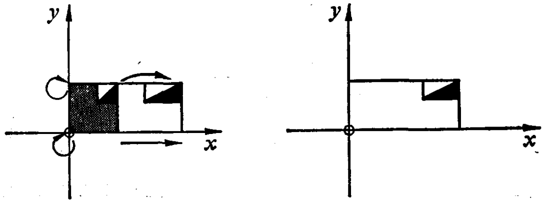
Original plane

Transformed plane

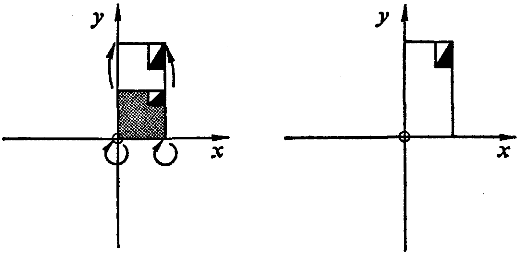
Positive and negative values of k give enlargements of the triangle as shown

The typical matrix for an enlargement is $\begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$ where $k \in R$.

Associated transformations



The matrix for this transformation would be $\begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix}$, $k \in R$.

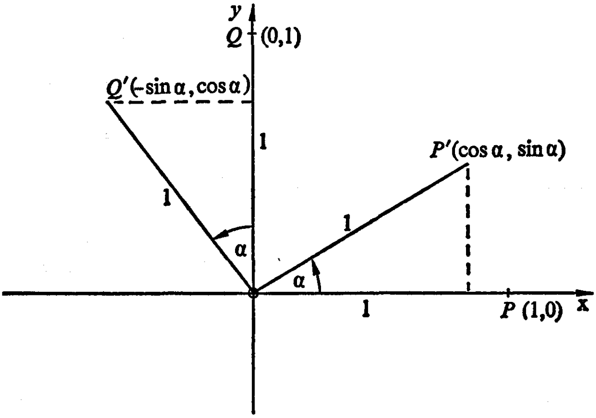


$\begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix}$

Negative values of k will, of course, give a transformed picture which will be a rotation together with an elongation.

4.7 (*b*) *Rotations*

To find the matrix which represents an anticlockwise rotation of $\alpha°$ about O.



We note that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ c \end{pmatrix}$$

and

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} b \\ d \end{pmatrix}$$

Hence we find the 'images' of the points $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ under this rotation. They are $\begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}$ and $\begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix}$. The required matrix is then

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

This can be used to give the formula of cos $(\alpha + \beta)$ for example since a rotation through $\alpha + \beta$ is a rotation through α followed by a rotation through β, i.e.

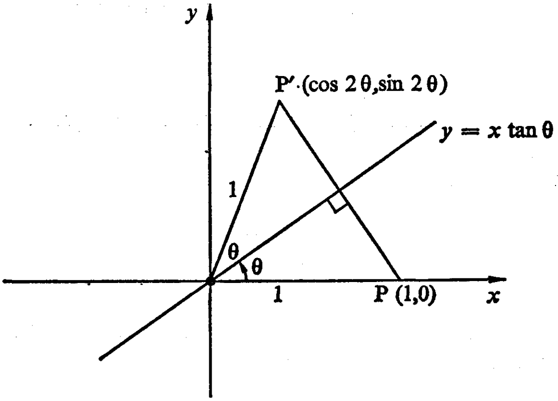
$$\begin{pmatrix}\cos(\alpha+\beta) & -\sin(\alpha+\beta)\\ \sin(\alpha+\beta) & \cos(\alpha+\beta)\end{pmatrix} = \begin{pmatrix}\cos\beta & -\sin\beta\\ \sin\beta & \cos\beta\end{pmatrix}\begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix}$$

and comparing entries

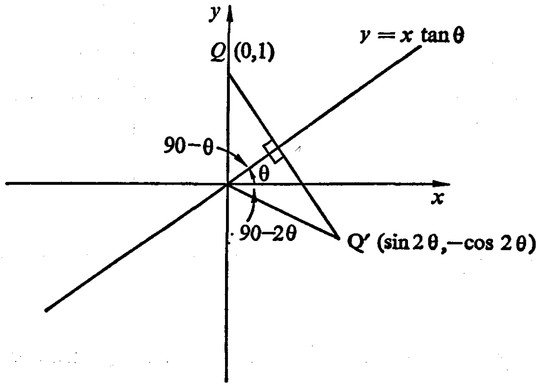
$$\cos(\alpha+\beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta$$
$$\sin(\alpha+\beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta$$

4.8 (*c*) *Reflections*

To find the matrix which represents a reflection in the line $y = x\tan\theta$. The image of (1, 0) is $(\cos 2\theta, \sin 2\theta)$.



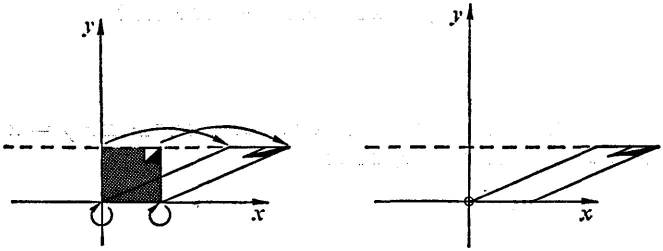
The image of (0, 1) is $(\sin 2\theta, -\cos 2\theta)$.



The matrix representing this reflection is thus

$$\begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix}$$

4.9 (*d*) *Shears*

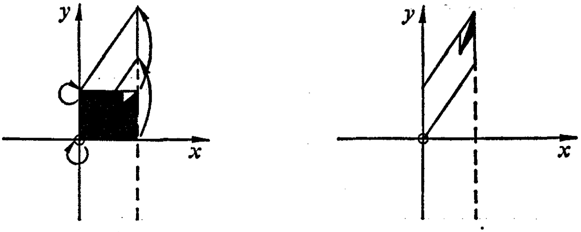


A shear in the x direction (which may be to the right or to the left depending on k).

A typical matrix is

$$\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$$

Note that the determinant must be 1. (Parallelograms on the same base and between the same parallels are equal in area.)



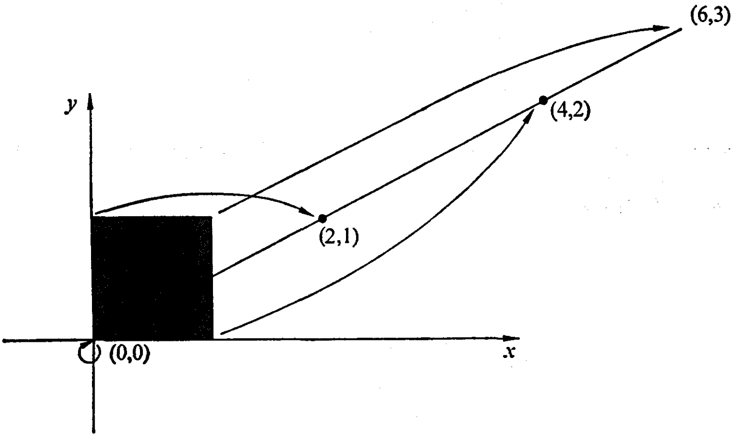
A shear in the y direction (which may be up or down, depending on the sign of k).

$\begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}$ is a typical matrix of the transformation, where $k \in R$.

4.10 (*e*) *Transformation involving a singular matrix*

Consider, for example, the transformation of the unit square by the matrix

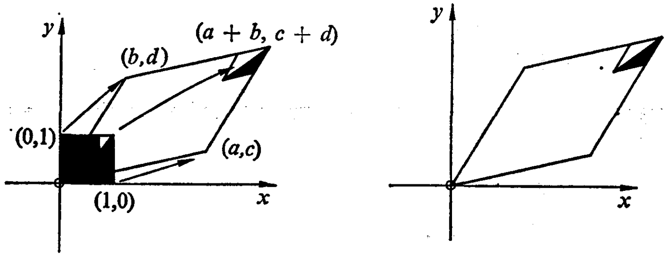
$$\begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix}$$



The unit square is transformed into part of the line $y = \frac{1}{2}x$.

4.11 (*f*) *Transformation by a general 2 × 2 matrix*

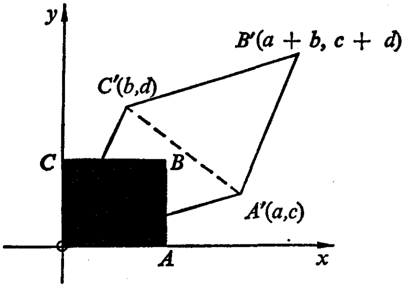
$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$



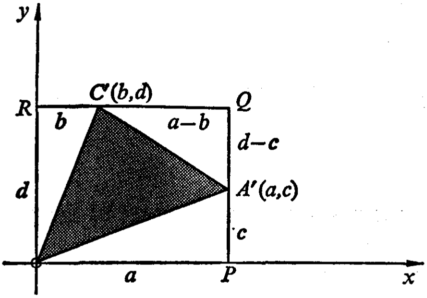
In the above example, it has been assumed that $ad - bc > 0$. What happens otherwise?

Determinants and the transformations

4.12 As above, the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ transforms the unit square in general, into a parallelogram.

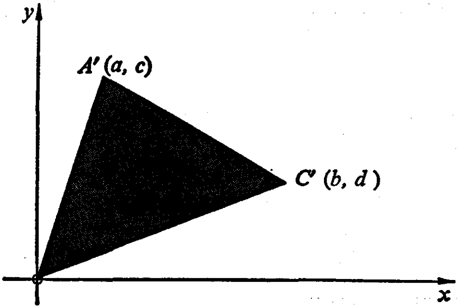


We shall interest ourselves in the ratio of the areas $OABC : OA'B'C'$. Now the area of $OABC$ is one square unit and the area of $OA'B'C'$ is twice the area of $\Delta OA'C'$.



$$\begin{aligned}
\text{Now} \qquad \Delta OA'C' &= OPQR - (\Delta OAP' + \Delta A'QC' + \Delta C'RO) \\
&= ad - \tfrac{1}{2}(ac + (a - b)(d - c) + bd) \\
&= \tfrac{1}{2}(ad - bc) \\
\text{So} \qquad OA'B'C' &= ad - bc
\end{aligned}$$

If the transformed figure is not as shown, but as below, then the area is $\tfrac{1}{2}(bc - ad)$.



Hence the area of the transformed figure is $\pm(ad - bc)$. The number $\Delta = ad - bc$ is the determinant (see **1.4**) of the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is written $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$. This number determines the 'area magnification' of the transformation, as opposed to linear magnification.

Successive transformations

4.13 Suppose two successive transformations are made so that (x, y) is transformed into (x', y'), where

$$\left.\begin{aligned} ax + by &= x' \\ cx + dy &= y' \end{aligned}\right\} \qquad (1)$$

$$\Rightarrow \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x' \\ y' \end{pmatrix} \qquad (2)$$

and then (x', y') is transformed into (x'', y'') where

$$\left.\begin{aligned} px' + qy' &= x'' \\ rx' + sy' &= y'' \end{aligned}\right\} \qquad (3)$$

$$\Rightarrow \begin{pmatrix} p & q \\ r & s \end{pmatrix}\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x'' \\ y'' \end{pmatrix} \qquad (4)$$

Substituting we have ((1) into (3))

$$\left.\begin{aligned} p(ax + by) + q(cx + dy) &= x'' \\ r(ax + by) + s(cx + dy) &= y'' \end{aligned}\right\} \qquad (5)$$

and also

$$\begin{pmatrix} p & q \\ r & s \end{pmatrix}\left[\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}\right] = \begin{pmatrix} x'' \\ y'' \end{pmatrix} \qquad (6)$$

from (2) and (4).

(5) gives

$$\left.\begin{aligned} (pa + qc)x + (pb + qd)y &= x'' \\ (ra + sc)x + (rb + sd)y &= y'' \end{aligned}\right\}$$

or

$$\begin{pmatrix} pa + qc & pb + qd \\ ra + sc & rb + sd \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x'' \\ y'' \end{pmatrix}$$

Comparing this with equation (6) we have

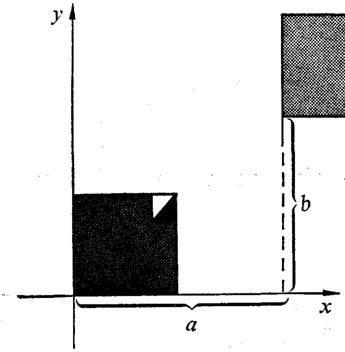
$$\begin{pmatrix} p & q \\ r & s \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} pa + qc & pb + qd \\ ra + sc & rb + sd \end{pmatrix}$$

which is our rule for multiplying 2 × 2 matrices.*

Transformations of the type given in equation (1) are called *linear.*

Translations and transformations with 3 × 3 matrices

4.14 It is not possible to transform $\begin{pmatrix} x \\ y \end{pmatrix}$ into $\begin{pmatrix} x + a \\ y + b \end{pmatrix}$ using 2 × 2 matrices. This transformation represents a translation a units in the x direction, b units in the y direction.



A translation $T_{(a,\ b)}$ of the unit square.

This translation can be represented in a number of ways

$$\left[\text{e.g. } \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} x + a \\ y + b \end{pmatrix}\right]^{\dagger}$$

With 3 × 3 matrices however, it is easy to achieve since

$$\begin{pmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} x + a \\ y + b \\ 1 \end{pmatrix}$$

Hence

$$\begin{pmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix}$$

represents a translation through a units in the x direction and through b units in the y direction.

---

* Hence it is possible to justify the rule for multiplication of matrices by starting with a definition of a linear transformation.

† See the section on Affine Geometry, p. 223.

To enable us to combine translations, rotations etc., we can easily extend our 2×2 matrices into 3×3 ones. For example, a rotation through an angle θ anticlockwise which is represented by

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad \text{becomes} \quad \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

A reflection in the x axis, represented by

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{becomes} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and so on. We then find the image of $\begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$ representing the point with coordinates (x, y).

Combining transformations

4.15 *Example 1.* To find the matrix representing a reflection in the line $y = x \tan\alpha$. The idea is to apply successive simple transformations to achieve the required result. Hence

Apply $R_{-\alpha}$, a rotation through an angle $-\alpha$ $\begin{pmatrix} \cos\alpha & +\sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}$

then X, a reflection in the x axis $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

and finally R_α, a rotation through an angle α $\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$

The matrix representing a reflection in $y = x \tan\alpha$ is given by

$$R_\alpha X R_{-\alpha} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}^*$$

$$= \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$$

$$= \begin{pmatrix} \cos 2\alpha & \sin 2\alpha \\ \sin 2\alpha & -\cos 2\alpha \end{pmatrix}$$

---

* Remember matrix multiplication is associative, hence there is no ambiguity in writing ABC for the product of the matrices A, B and C.

4.16 *Example 2.* To find the matrix representing a reflection in the line $y = mx + c$.

Suppose that $m = \tan\alpha$. As before

Apply

$T_{(0,-c)}$, a translation of $-c$ in the y direction $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -c \\ 0 & 0 & 1 \end{pmatrix}$

then

$R_{-\alpha}$, a rotation through an angle of $-\alpha$ $\begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}$

X, a reflection in the x axis $\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

R_α, a rotation of through an angle α $\begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}$

and finally

$T_{(o,\,c)}$, a translation of $+c$ in the y direction $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}$

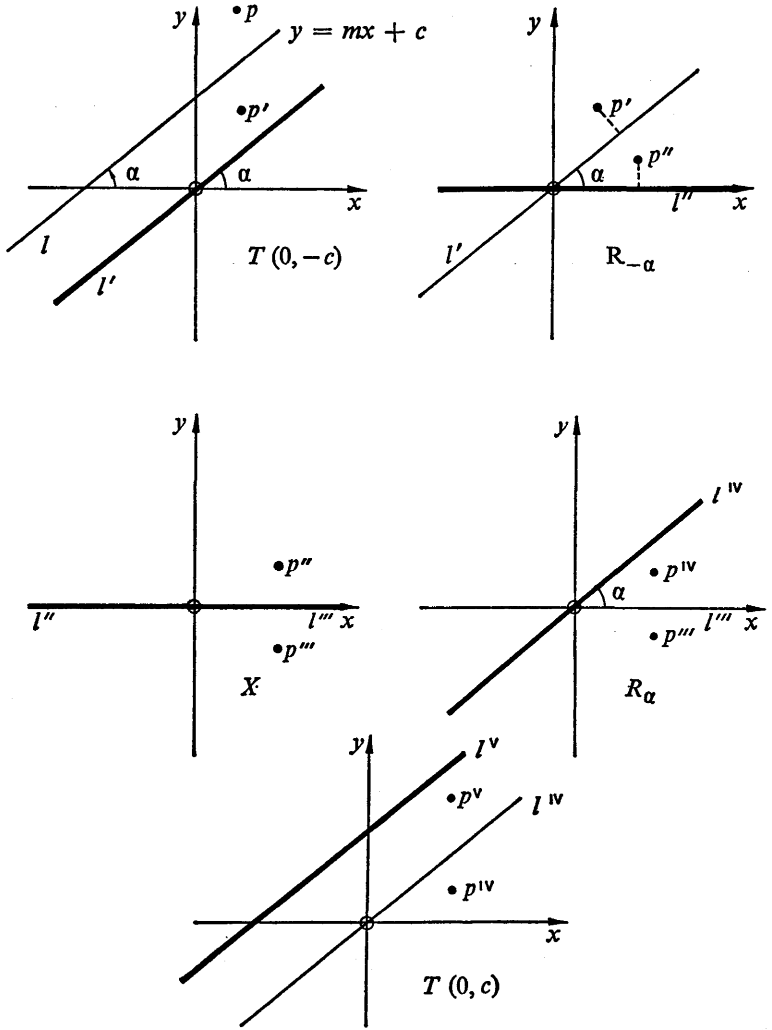
The required matrix is thus $T_{(o,\,c)}R_\alpha XR_{-\alpha}T_{(o,\,-c)}$[*] which gives

$$\begin{pmatrix} \cos 2\theta & \sin 2\theta & -c\sin 2\theta \\ \sin 2\theta & -\cos 2\theta & c(\cos 2\theta + 1) \\ 0 & 0 & 1 \end{pmatrix},$$

on multiplication. If a solution involving m is required, this can easily be found since if $m = \tan\theta$, then $\dfrac{m}{\sqrt{1+m^2}} = \sin\theta$ and $\dfrac{1}{\sqrt{1+m^2}} = \cos\theta$.

We can illustrate this as shown.

[*] Remember matrix multiplication is associative, hence there is no ambiguity in writing ABC for the product of the matrices A,B,C.



Transformations of three-dimensional space using matrices

4.17 To investigate transformations of 3*d* space, we proceed, as with 2*d* space except that this time we consider the transformation of the 'unit cube' with vertices (0, 0, 0), (1, 0, 0), (0, 1, 0), (0, 0, 1), (1, 1, 0), (1, 0, 1), (0, 1, 1) and (1, 1, 1).

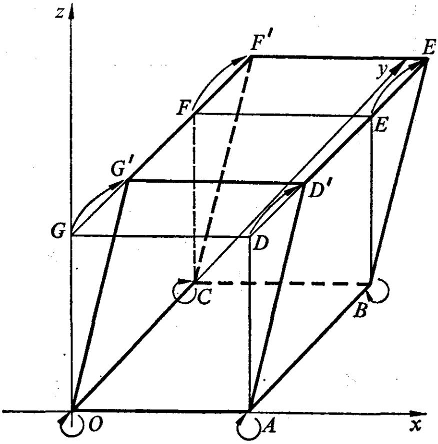
Note again that translations cause difficulty and we would need to consider 4 × 4 matrices with (x, y, z) represented by $\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$.

Example

Investigate the transformation of the unit cube represented by the matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix}$$

O	(0, 0, 0)	transforms into	(0, 0, 0)	O
A	(1, 0, 0)	transforms into	(1, 0, 0)	A
C	(0, 1, 0)	transforms into	(0, 1, 0)	C
G	(0, 0, 1)	transforms into	$(0, \frac{1}{2}, 1)$	G'
B	(1, 1, 0)	transforms into	(1, 1, 0)	B
D	(1, 0, 1)	transforms into	$(1, \frac{1}{2}, 1)$	D'
F	(0, 1, 1)	transforms into	$(0, \frac{3}{2}, 1)$	F'
E	(1, 1, 1)	transforms into	$(1, \frac{3}{2}, 1)$	E'



Orthogonal* transformations

4.18 *Definition*

A square matrix A is said to be *orthogonal* if

$$A'A = I \quad \text{or} \quad A' = A^{-1}$$

---

* We assume some knowledge of vectors here, i.e.

(i) Scalar product of two vectors $\left[\begin{pmatrix}\alpha\\ \beta\end{pmatrix}\right.$ and $\begin{pmatrix}\gamma\\ \delta\end{pmatrix}$ is $(\alpha\ \beta)\begin{pmatrix}\gamma\\ \delta\end{pmatrix} = \alpha\gamma + \beta\delta\Big]$.

(ii) $\cos\theta = \dfrac{\mathbf{a}\,.\,\mathbf{b}}{|\mathbf{a}|\,|\mathbf{b}|}$.

(iii) Orthogonality when $\theta = 90°$, i.e. $\mathbf{a}\,.\,\mathbf{b} = 0$.

[A' is the transposed matrix of A, i.e. rows and columns of A are interchanged.] Since

$$AA^{-1} = A^{-1}A = I$$

then

$$AA' = A'A = I$$

Typical examples of orthogonal matrices are

$$\begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & -\frac{3}{5} \end{pmatrix}, \quad \begin{pmatrix} \dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} & -\dfrac{1}{\sqrt{2}} \end{pmatrix}, \quad \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{pmatrix}$$

since e.g.

$$\begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & -\frac{3}{5} \end{pmatrix}\begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & -\frac{3}{5} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and

$$\begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{pmatrix}\begin{pmatrix} \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

4.19 *Properties*

(*a*) If A and B are orthogonal we have

$$AA' = A'A = I \quad \text{and} \quad BB' = B'B = I$$

Hence

$$\begin{aligned} (AB)(AB)' &= AB\,.\,B'A' \qquad \text{(since } (AB)' = B'A'\text{)} \\ &= A(BB')A' \\ &= AIA' \\ &= AA' \\ &= I \end{aligned}$$

Hence the product of two orthogonal matrices is another orthogonal matrix.

(*b*) The inverse of an orthogonal matrix is orthogonal and also the transpose of an orthogonal matrix is orthogonal. Prove these for yourself.

(*c*) The determinant of an orthogonal matrix is ±1, since $AA' = I$,

$$|A||A'| = |A|^2 = |I| = 1$$

Hence

$$|A| = \pm 1$$

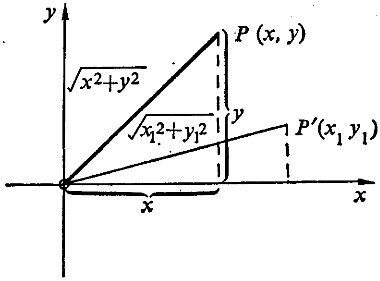
(*d*) Consider a point P (x, y) in a plane. Call the column vector

$$\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{X}.$$

Then

$$X'X = (x \quad y)\begin{pmatrix} x \\ y \end{pmatrix} = x^2 + y^2$$

which is the square of the distance of P from the origin.



Apply an orthogonal transformation A to X. Our new co-ordinates $(x_1 \quad y_1)$ $\left(\text{with associated column vector } \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}\right)$ = Y, are given by

$$Y = AX$$

since $Y'Y = x_1^2 + y_1^2$ and O remains the same point,

$$\begin{aligned} Y'Y &= (AX)' \,.\, AX \\ &= X'A' \,.\, AX \\ &= X'X \end{aligned}$$

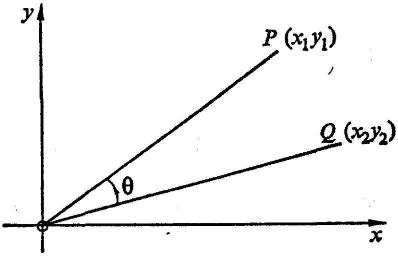
i.e.

$$x_1^2 + y_1^2 = x^2 + y^2$$

Hence under an orthogonal transformation, distance is preserved.

(*e*) Another important feature of this type of transformation is that the angle between two vectors is invariant (i.e. remains the same).

Consider



where

$$X = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$$

$$Y = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$$

Then $$\cos\theta = \frac{x_1x_2 + y_1y_2}{\sqrt{x_1^2 + y_1^2}\,.\,\sqrt{x_2^2 + y_2^2}}$$

which is also given by

$$\cos\theta = \frac{\mathbf{X}'\mathbf{Y}}{\{\mathbf{X}'\mathbf{X}\}^{\frac{1}{2}}\{\mathbf{Y}'\mathbf{Y}\}^{\frac{1}{2}}}$$

Suppose we now apply an orthogonal transformation A to each of $\mathbf{X}$ and $\mathbf{Y}$ so that

$$\mathbf{Z} = A\mathbf{X}, \quad \text{where} \quad \mathbf{Z} = \begin{pmatrix} z_1 \\ t_1 \end{pmatrix}$$

$$\mathbf{T} = A\mathbf{Y}, \quad \text{where} \quad \mathbf{T} = \begin{pmatrix} z_2 \\ t_2 \end{pmatrix}$$

Now
$$\begin{aligned}
\cos\phi &= \frac{\mathbf{Z}'\,.\,\mathbf{T}}{\{\mathbf{Z}'\mathbf{Z}\}^{\frac{1}{2}}\{\mathbf{T}'\mathbf{T}\}^{\frac{1}{2}}} \\
&= \frac{(A\mathbf{X})'\,.\,A\mathbf{Y}}{\{(A\mathbf{X})'(A\mathbf{X})\}^{\frac{1}{2}}\{(A\mathbf{Y})'\,.\,(A\mathbf{Y})\}^{\frac{1}{2}}} \\
&= \frac{\mathbf{X}'A'A\mathbf{Y}}{\{\mathbf{X}'A'A\mathbf{X}\}^{\frac{1}{2}}\{\mathbf{Y}'A'A\mathbf{Y}\}^{\frac{1}{2}}} \\
&= \frac{\mathbf{X}'\mathbf{Y}}{\{\mathbf{X}'\mathbf{X}\}^{\frac{1}{2}}\{\mathbf{Y}'\mathbf{Y}\}^{\frac{1}{2}}} \qquad \text{because } A'A = I. \\
&= \cos\theta.
\end{aligned}$$

Hence $\theta = \phi$* and the angle is preserved by an orthogonal transformation.

General conditions for a 2 × 2 matrix to be orthogonal

4.20 Suppose

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

then we require $A'A = I$, or

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

so, comparing entries:

$$a^2 + c^2 = 1, \quad ab + cd = 0, \quad b^2 + d^2 = 1$$

If $\quad a \in R$ and $c \in R$

then $\quad a^2 \geqslant 0$ and $c^2 \geqslant 0$

---

* For justification, read on.

Hence if $a^2 + c^2 = 1$,

$$a^2 \leqslant 1 \quad \text{and} \quad c^2 \leqslant 1$$

$$\Rightarrow -1 \leqslant a \leqslant 1 \quad \text{and} \quad -1 \leqslant c \leqslant 1$$

Similarly for b and d.

Hence it is in order to let $a = \cos\theta$, $b = \cos\phi$, $c = \sin\theta$, $d = \sin\phi$ and we have

$$\cos\theta\cos\phi + \sin\theta\sin\phi = 0$$

i.e.

$$\cos(\theta - \phi) = 0$$

$$\Rightarrow \theta - \phi = \frac{\pi}{2} \quad \text{or} \quad \theta - \phi = \frac{3\pi}{2}$$

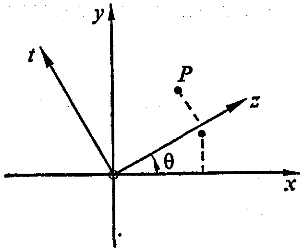
There are only two 2×2 matrices which satisfy our conditions

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$$

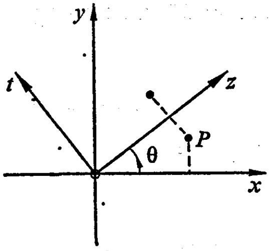
The first matrix represents a rotation of $\theta°$ about the origin, whilst the second corresponds to a rotation of $\theta°$ followed by a reflection in the new x-axis (i.e. z axis as shown).

In the first case the determinant is 1, in the second -1.

Case 1



Case 2



Quadratic form

4.21 We define a quadratic form $Q(x, y)$ as

$$Q(x, y) = ax^2 + 2hxy + by^2$$

Let

$$\mathbf{X} = \begin{pmatrix} x \\ y \end{pmatrix}, \quad A = \begin{pmatrix} a & h \\ h & b \end{pmatrix}$$

Then

$$\mathbf{X}' = (x \quad y)$$

and
$$
\begin{aligned}
X'AX &= (x \quad y)\begin{pmatrix} a & h \\ h & b \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \\
&= (x \quad y)\begin{pmatrix} ax + hy \\ hx + by \end{pmatrix} \\
&= ax^2 + 2hxy + by^2 \\
&= Q(x, y)
\end{aligned}
$$

A is called the matrix of the quadratic form $Q(x, y)$. Note that A is symmetric (i.e. $A' = A$). The eigenvectors of a symmetric matrix are orthogonal (see below).

4.22 *Example.* Find the eigenvectors of the matrix

$$A = \begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix}$$

and show that they are orthogonal.

Eigenvectors were mentioned on p. 85. They are those whose direction is unchanged by the matrix transformation A, i.e.

$$\begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix}$$

$$\Rightarrow \left.\begin{aligned} x + 2y &= \lambda x \\ 2x - 2y &= \lambda y \end{aligned}\right\}$$

$$\Rightarrow \left.\begin{aligned} x(1 - \lambda) &= -2y \\ 2x &= (\lambda + 2)y \end{aligned}\right\}$$ and since we do not require the solution $x = 0, y = 0$

$$
\begin{aligned}
&\Rightarrow \frac{1 - \lambda}{2} = \frac{-2}{\lambda + 2} \\
&\Rightarrow (1 - \lambda)(\lambda + 2) = -4 \\
&\Rightarrow \lambda^2 + \lambda - 6 = 0 \\
&\Rightarrow \lambda = 2 \quad \text{or} \quad \lambda = -3
\end{aligned}
$$

When $\lambda = 2$
$$
\begin{aligned}
x + 2y &= 2x \\
\Rightarrow x &= 2y.
\end{aligned}
$$

Any vector satisfying this relation is suitable, e.g. $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$. When $\lambda = -3$,

$$x + 2y = -3x$$
i.e. $$-\tfrac{1}{2}y = x$$

A suitable vector is

$$\begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

The scalar product of these two vectors is given by

$$(2\quad 1)\begin{pmatrix}1\\-2\end{pmatrix}=0$$

Hence they are orthogonal.

4.23 *Prove* that this is true in general, i.e. the eigenvectors of a symmetric matrix (with unequal eigenvalues (values of λ)) are orthogonal. Consider only the 2 × 2 matrix

$$A=\begin{pmatrix}a & h\\ h & b\end{pmatrix}$$

Orthogonal matrices have already be defined. If we consider the columns of an orthogonal matrix as column vectors, then we notice that these column vectors are of unit length and are mutually orthogonal.

Example. For $\begin{pmatrix}\frac{3}{5} & \frac{4}{5}\\ \frac{4}{5} & -\frac{3}{5}\end{pmatrix}$; we have

$$(\tfrac{3}{5}\quad \tfrac{4}{5})\begin{pmatrix}\frac{3}{5}\\ \frac{4}{5}\end{pmatrix}=1 \quad\text{and}\quad (\tfrac{4}{5}\quad -\tfrac{3}{5})\begin{pmatrix}\frac{4}{5}\\ -\frac{3}{5}\end{pmatrix}=1$$

Hence each column vector is of unit length. Also

$$(\tfrac{3}{5}\quad \tfrac{4}{5})\begin{pmatrix}\frac{4}{5}\\ -\frac{3}{5}\end{pmatrix}=0$$

and

$$(\tfrac{4}{5}\quad -\tfrac{3}{5})\begin{pmatrix}\frac{3}{5}\\ \frac{4}{5}\end{pmatrix}=0$$

Hence the column vectors are mutually orthogonal.

Now if the eigenvectors are

$$\begin{pmatrix}2\\1\end{pmatrix}\quad\text{and}\quad\begin{pmatrix}1\\-2\end{pmatrix}$$

we can then make them of unit length by using instead

$$\begin{pmatrix}\dfrac{2}{\sqrt{2^2+1^2}}\\ \dfrac{1}{\sqrt{2^2+1^2}}\end{pmatrix}\quad\text{and}\quad\begin{pmatrix}\dfrac{1}{\sqrt{1^2+(-2)^2}}\\ \dfrac{-2}{\sqrt{1^2+(-2)^2}}\end{pmatrix}$$

i.e.

$$\begin{pmatrix}\dfrac{2}{\sqrt{5}}\\ \dfrac{1}{\sqrt{5}}\end{pmatrix}\quad\text{and}\quad\begin{pmatrix}\dfrac{1}{\sqrt{5}}\\ \dfrac{-2}{\sqrt{5}}\end{pmatrix}$$

Now consider the matrix

$$V = \begin{pmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{5}} \end{pmatrix}$$

Clearly $V'V = I$, and so V is orthogonal.

4.24 *Prove* that it is possible to write

$$AV = VD \quad \text{where} \quad A = \begin{pmatrix} a & h \\ h & b \end{pmatrix}$$

V is an orthogonal matrix as above, and D is the diagonal matrix

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

where λ_1 and λ_2 are the eigenvalues.

A transformation of (x, y) into (x_1, y_1) by the transformation $\mathbf{X} = T\mathbf{X}_1$ transforms a quadratic form with matrix A into one with matrix $T'AT$. Now

$$\mathbf{X}'A\mathbf{X} = Q(x, y)$$

Substituting

$$(T\mathbf{X}_1)'AT\mathbf{X}_1 = \mathbf{X}_1'T'AT\mathbf{X}_1$$

which is another quadratic form of matrix $T'AT$. Suppose we can choose T so that $T'AT$ *is* a diagonal matrix E with

$$E = \begin{pmatrix} e_1 & 0 \\ 0 & e_2 \end{pmatrix}$$

and so

$$Q(x, y) = \mathbf{X}_1' \begin{pmatrix} e_1 & 0 \\ 0 & y_1 \end{pmatrix} \mathbf{X}_1$$

$$= e_1(x_1)^2 + e_2(x_2)^2$$

Therefore we choose $T = V$, then

$$T'AT = V'AV \quad \text{and} \quad V'AV = D^*$$

also $\quad \mathbf{X} = T\mathbf{X}_1 \quad \text{or} \quad \mathbf{X} = V\mathbf{X}_1$

so $\quad \mathbf{X}_1 = V'\mathbf{X}$

Hence a linear transformation $\mathbf{X}_1 = V'\mathbf{X}$ produces a new quadratic form $\lambda_1(x_1)^2 + \lambda_2(y_1)^2$, where λ_1, λ_2 are the eigenvalues of A.

---

* See previous page.

4.25 This is most useful in the investigation of conics. The application of the linear transformation $\mathbf{X}_1 = V'\mathbf{X}$ transforms the equation of a conic $\mathbf{X}'A\mathbf{X} = k$ into a conic of the form $\lambda_1(x_1)^2 + \lambda_2(x_2)^2 = k$ provided that V represents a rotation (see p. 86) (i.e. V is not symmetric).

Example

Investigate the conic $x^2 - 4xy - 2y^2 = 3$,

$$Q(x, y) \equiv x^2 - 4xy - 2y^2$$
$$\equiv (x \quad y)\begin{pmatrix} 1 & -2 \\ -2 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

Here $$A = \begin{pmatrix} 1 & -2 \\ -2 & -2 \end{pmatrix}$$

The eigenvalues are

$$\lambda = 2 \quad \text{and} \quad \lambda = -3$$

and so the conic is transformed into

$$2x_1^2 - 3y_1^2 = 3$$

by the application of the transformation $\mathbf{X}_1 = V'\mathbf{X}$.

V is easily found as

$$\begin{pmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix}$$

(remember V must not be symmetric), and

$$V' = \begin{pmatrix} \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix}$$

Finally, the transformation is given by

$$\mathbf{X}_1 = V'\mathbf{X}$$

i.e. $$x_1 = \frac{2}{\sqrt{5}}x - \frac{1}{\sqrt{5}}y$$

$$y_1 = \frac{1}{\sqrt{5}}x + \frac{2}{\sqrt{5}}y$$

This represents a rotation through the angle $\cos^{-1}\frac{2}{\sqrt{5}}$.

Hence the conic

$$x^2 - 4xy - 2y^2 = 3$$

can be transformed, by rotating the axes, into

$$2x_1^2 - 3y_1^2 = 3$$

The new conic is

$$\frac{x_1^2}{\frac{3}{2}} - \frac{y_1^2}{1} = 1$$

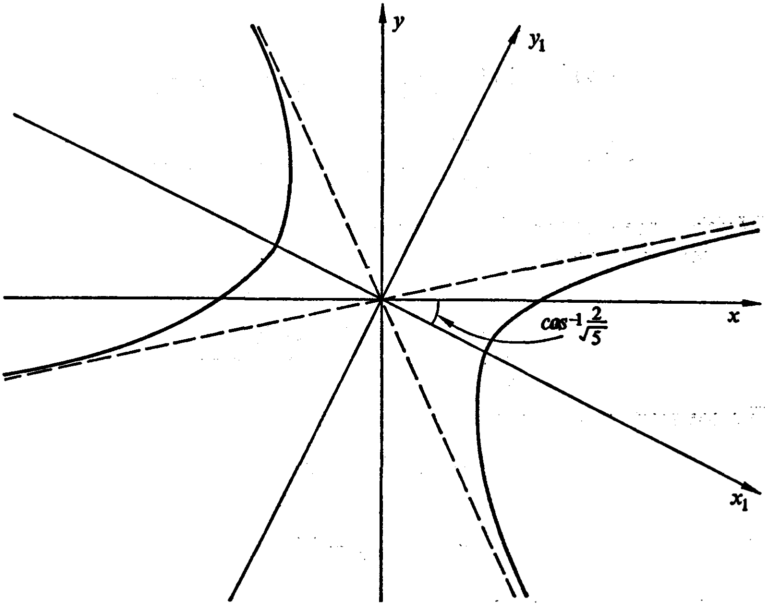
which is of the form

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

and so represents an hyperbola with centre the origin, and asymptotes

$$x_1 = \pm\sqrt{\tfrac{3}{2}}\,y_1$$

We may sketch the curve thus



4.26 The non-homogeneous quadratic form is given by

$$Q(x, y) = ax^2 + 2hxy + by^2 + 2gx + 2fy + c$$

This can be written in the form

$$\mathbf{X}'A\mathbf{X} + B\mathbf{X} + C, \qquad \text{where } B = (2g\ 2f)$$

Application of a rotation by an orthogonal matrix, followed by completion of the square will reduce these conics to a standard form.

Example

Investigate the conic

$$x^2 - 4xy - 2y^2 + 6x - 4y + 3\tfrac{1}{3} = 0$$

As above this becomes,

$$(xy)\begin{pmatrix} 1 & -2 \\ -2 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + (6 \quad -4)\begin{pmatrix} x \\ y \end{pmatrix} + 3\tfrac{1}{3} = 0$$

We now apply $\mathbf{X}_1 = V'\mathbf{X}$ which transforms the conic to be

$$2x_1^2 - 3y_1^2 + (6 - 4)\begin{pmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix}\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + 3\tfrac{1}{3} = 0$$

$$\Rightarrow 2x_1^2 - 3y_1^2 + \frac{16}{\sqrt{5}}x_1 - \frac{2}{\sqrt{5}}y_1 + 3\tfrac{1}{3} = 0$$

$$\Rightarrow 2\left(x_1 + \frac{4}{\sqrt{5}}\right)^2 - 3\left(y_1 + \frac{1}{3\sqrt{5}}\right)^2 - 3 = 0$$

Which is the standard form for a conic.

It can be written as

$$\frac{\left(x_1 + \frac{4}{\sqrt{5}}\right)^2}{(\frac{3}{2})^2} - \frac{\left(y_1 + \frac{1}{3\sqrt{5}}\right)^2}{1^2} = 1$$

We see from this, that if

$$\left.\begin{aligned} x_1 + \frac{4}{\sqrt{5}} &= \mathrm{X} \\ y_1 + \frac{1}{3\sqrt{5}} &= \mathrm{Y} \end{aligned}\right\}$$

then the conic is in the form

$$\frac{\mathrm{X}^2}{a^2} - \frac{\mathrm{Y}^2}{b^2} = 1$$

The transformations

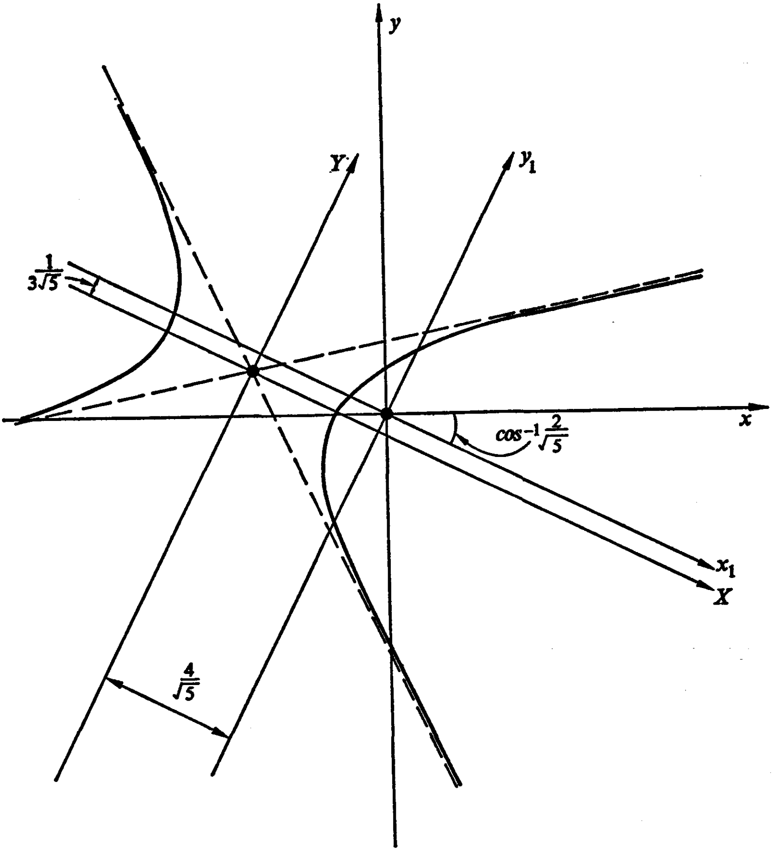
$$x_1 + \frac{4}{\sqrt{5}} = X, \quad y_1 + \frac{1}{3\sqrt{5}} = Y$$

are equivalent to a translation.

Hence the conic is an hyperbola with centre

$$\left(-\frac{4}{\sqrt{5}}, -\frac{1}{3\sqrt{5}}\right)$$

It can be drawn graphically as follows, as in the simpler example above.



Exercise 4.1

1. Prove that rotation through a right angle about the origin is given by the transformation matrix

$$T = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

The point $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ is subjected to the transformation T, followed by the further transformation

$$T_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Show that the result is equivalent to a reflection in the line $y + x = 0$. Find the equation of the line resulting from the transformations T_1 and T on the line $y = 3x + 2$. (*MEI Specimen*)

2. Show that the 2×2 matrix M_α which reflects each point of the xy plane in the line $y = x \tan \alpha$ is

$$M_\alpha = \begin{pmatrix} \cos 2\alpha & \sin 2\alpha \\ \sin 2\alpha & -\cos 2\alpha \end{pmatrix}$$

If M_β represents a reflection in the line $y = x \tan \beta$, show that $M_\beta M)$ represents a rotation of the plane, and find the angle of rotation. Find the relation satisfied by α and β if $M_\beta M_\alpha = M_\alpha M_\beta$. (*MEI Specimen*)

3. Describe how complex numbers may be represented by 2×2 matrices. Find A^6 where

$$A = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}$$

(*Part MEI Pure I*)

4. If

$$M = \begin{pmatrix} 1 & -1 & k \\ 4 & 7 & 3 \\ 3 & 19 & 1 \end{pmatrix}$$

find det M in terms of k.

(i) If $k = 2$, solve

$$M\mathbf{x} = \mathbf{O} \quad \text{and} \quad M\mathbf{x} = \begin{pmatrix} 3 \\ 8 \\ 16 \end{pmatrix}$$

(ii) If $k = 1$, solve $M\mathbf{x} = \mathbf{O}$ and interpret your result geometrically. (*MEI Mathematics I*)

5. If M denotes the matrix $\begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$ and I denotes the 2×2 unit matrix, prove that $M^2 = 3M - 2I$.

* [See *Matrices 2* Chapter 2.]

Prove further that, if n is any positive integer

$$M^n = (2^n - 1)M - 2(2^{n-1} - 1)I \qquad (MEI\ Pure\ I)$$

6. (*a*) If

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 4 & -5 \\ 3 & 7 & -9 \end{pmatrix}$$

Find $|A|$ and A^{-1}.

(*b*) Write the transformations

$$\left.\begin{aligned} X &= x + 2y - z \\ Y &= 2x + 4y - 5z \\ Z &= 3x + 7y - 9z \end{aligned}\right\} \qquad \left.\begin{aligned} X &= a + b - c \\ Y &= 3a - 2b + c \\ Z &= 4a + 3b + 2c \end{aligned}\right\}$$

in matrix form and hence, by using matrices, express x, y, z

(i) in terms of X, Y, Z, (ii) in terms of a, b, c.

7. If B and C are square unequal matrices of the same order and I is the unit matrix, why does

$$(I + B)(I - B) = (I - B)(I + B)$$

whereas $(I + B)(I - C)$ is not in general equal to $(I - C)(I + B)$? If

$$A = \begin{pmatrix} 1 & c & -b \\ -c & 1 & a \\ b & -a & 1 \end{pmatrix} \quad \text{and} \quad B = I - A$$

show that

$$AA' = (I + B)(I - B) = A'A$$

where A' is the transpose of A.

8. Investigate the transformation of the unit cube by the matrix

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

and show that the volume of the transformed solid is given by

$$\pm \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}$$

9. Investigate the conics:

(i) $9xy - 1 = 0$,
(ii) $3x^2 + 8xy + 9y^2 + x - 2y + 5 = 0$.

10. Investigate the conics:

(i) $x^2 + 4xy + 4y^2 + 2x + 4y + 1 = 0$,
(ii) $x^2 + 2y^2 - 2x + 4y - 1 = 0$.

11. If a rotation of the coordinate axes about the origin transforms the equation

$$ax^2 + 2hxy + by^2 = 1$$

into the equation

$$a'x'^2 + 2h'x'y' + b'y'^2 = 1,$$

prove that

$$a' + b' = a + b \quad \text{and} \quad a'b' - h'^2 = ab - h^2 \qquad (O \,\&\, C)$$

12. If the equation of the parabola

$$(x \sin \theta + y \cos \theta)^2 + 2gx + 2fy + c = 0$$

is reduced to the standard form $Y^2 = kX$ by a change of coordinate axes (leaving the unit of length unaltered), prove that

$$k = 2(f \sin \theta - g \cos \theta)$$

Hence, or otherwise, find the distance between the focus and the directrix of the parabola given by the equation

$$9x^2 + 24xy + 16y^2 - 10x + 20y = 0 \qquad (O \,\&\, C)$$

Exercise 4.2 Miscellaneous

1. Let $A(\theta)$ and $B(\theta)$ denote the matrices

$$\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}, \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

(i) If

$$\begin{pmatrix} X' \\ Y' \end{pmatrix} = A(\theta) \begin{pmatrix} X \\ Y \end{pmatrix}$$

show that the point (X', Y') is the mirror image of (X, Y) in the line $y = x \tan \dfrac{\theta}{2}$

(ii) Prove that $A(\theta_1)A(\theta_2) = B(\phi)$ where ϕ is to be determined and hence explain the relation between the points (X', Y') and (X, Y) when

$$\begin{pmatrix} X' \\ Y' \end{pmatrix} = B(\theta) \begin{pmatrix} X \\ Y \end{pmatrix}$$

(iii) Prove that

$$A(\theta_1)A(\theta_2)A(\theta_3) = A(\theta_3)A(\theta_2)A(\theta_1)$$

and interpret this result geometrically. (*CS*)

2. (i) Find a 3×3 matrix P such that, for any 3×3 matrix Q, PQ is the matrix Q with its second and third rows interchanged.

(ii) The direction cosines of a line in $3d$ Euclidean space are (l, m, n). Interpret geometrically the linear transformation

$$\mathbf{X} = L\mathbf{x}$$

where

$$L = \begin{pmatrix} l^2 & lm & ln \\ ml & m^2 & mn \\ nl & nm & n^2 \end{pmatrix}$$

Prove that the transformation L^2 is the same as the transformation L. (*MEI* '*S*')

3. The equations of motion of a particle in a plane, referred to rectangular axes Ox, Oy in the plane are $\ddot{x} = ky$, $\ddot{y} = -kx$. Show that the equations of motion become $\ddot{x}' = ky'$, $\ddot{y}' = -kx'$ and are thus unaltered in form when referred to axes Ox', Oy' derived from the first by rotation about O through an arbitrary angle α.

Are the equations of motion unaltered in form when referred to axes Ox'', Oy'' derived from Ox, Oy by reflection in an arbitrary line $\dfrac{x}{\cos\beta} = \dfrac{y}{\sin\beta}$ through the origin. Similarly investigate the equations:

$$\ddot{x} = k(x\dot{y} - y\dot{x})y, \quad \ddot{y} = -k(x\dot{y} - y\dot{x})x \qquad (CS)$$

4. If X' denotes the transpose of a matrix X and A and B are 3×3 matrices, prove that

$$(A + B)' = A' + B'$$

and

$$(A.B)' = B'A'$$

Deduce that the angle between any two vectors in $3d$ Euclidean space is unaltered when the space is subject to a transformation $\mathbf{Y} = A\mathbf{X}$ where A is a 3×3 matrix such that $A'.A = I$.* (*MEI mod.*)

5. Let A, B, C be real 2×2 matrices and write

$$[A, B] = AB - BA, \text{ etc.}$$

Prove that

(i) $[A, A] = 0$,
(ii) $[[A, B], C] + [[B, C], A] + [[C, A], B] = 0$,
(iii) $[A, B] = I \Rightarrow [A, B^m] = mB^{m-1}$ for all positive integers m.

---

* See section on orthogonal transformations.

At each step you should state clearly any properties of the matrix ring which you use.

The *trace* Tr (A) of a matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

is defined by

$$\operatorname{Tr}(A) = a_{11} + a_{22}$$

Prove that

(iv) $\operatorname{Tr}(A + B) = \operatorname{Tr}(A) + \operatorname{Tr}(B)$,
(v) $\operatorname{Tr}(AB) = \operatorname{Tr}(BA)$,
(vi) $\operatorname{Tr}(I) = 2$.

Deduce that there are no matrices satisfying $[A, B] = I$. Does this in any way invalidate the statement in (iii). *(MEI Pure III)*

6. The number $a_{11} + a_{22} + a_{33}$ is called the *trace* of the matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

If A and B are two 3×3 matrices, show that the traces of the matrices AB and BA are equal.

If the matrix AB represents a rotation through an angle ϕ about the directed axis U and A represents a rotation interchanging the axes U and V, explain why BA represents a rotation through the angle ϕ about V.

Given that the matrix

$$M = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

represents a rotation through the angle ϕ about the z-axis, and that the matrix C represents a rotation about *some* axis, find a formula for the angle of rotation in terms of the trace of C. *(CS)*

7. What is meant by the statement that multiplication of real numbers is associative and distributive over addition?

Prove that, if a_{ij}, b_{ij}, c_{ij} $(i, j = 1, 2, 3)$ are real numbers then

$$\sum_{\beta=1}^{3}\left(\sum_{\alpha=1}^{3} a_{i\alpha}b_{\alpha\beta}\right)c_{\beta j} = \sum_{\alpha=1}^{3} a_{i\alpha}\left(\sum_{\beta=1}^{3} b_{\alpha\beta}c_{\beta j}\right)$$

Deduce that multiplication of 3×3 matrices is associative and that, if A is a 3×3 real matrix, then the notation $A^3, A^4, A^5 \dots$ may be interpreted unambiguously. If the 3×3 real matrices B and $A^2 + A + I$ are non singular (I is the identity matrix) why should the notation $\frac{A}{B}$ be ambiguous but $\frac{A}{A^2 + A + I}$ not so? *(MEI Pure III)*

8. Let A be a 3×3 real matrix and $\mathbf{b}$ a 3 rowed column vector. It is proposed to solve the equation

$$A\mathbf{x} = \mathbf{b},$$

and a particular solution $\mathbf{x} = \mathbf{x}_o$ is noted. Prove that any other solution may be written in the form

$$\mathbf{x} = \mathbf{x}_o + \mathbf{u}$$

where $\mathbf{u}$ is a solution of the equation $A\mathbf{x} = \mathbf{O}$. Prove conversely that any vector $\mathbf{x}_o + \mathbf{u}$ is a solution of the equation where $\mathbf{x}_o$ is a fixed particular solution of the equation and u is any vector such that $A\mathbf{u} = \mathbf{O}$. Interpret geometrically the equation and its solutions if A is singular (i) when an $\mathbf{x}_o$ exists, and (ii) when there is no such particular solution $\mathbf{x}_o$. *(MEI Pure III)*

9. A function Δ maps the ring of real 2×2 matrices to the real numbers by assigning to any real 2×2 matrix A, its determinant ΔA. Prove

(i) $\Delta(AB) = \Delta A \,.\, \Delta B$ for all A, B.
(ii) $\Delta(A + B) \neq \Delta A + \Delta B$ in general.
(iii) Under matrix multiplication, the set of matrices A such that $\Delta A = 1$ forms a group, but under matrix multiplication and addition this set does not form a ring. *(SMP mod.)*

10. At time zero an observer at the origin O of one coordinate system, (x, y, z, t) coincides with an observer at the origin O_1, of another (x_1, y_1, z_1, t_1). At this time, $t = t_1 = 0$, an instantaneous light signal is emitted.

After time t the first observer describes the equation of the wavefront as

$$x^2 + y^2 + z^2 - c^2t^2 = 0 \tag{i}$$

where c is the velocity of light.

Similarly the second observer describes the equation at time t_1 as

$$x_1^2 + y_1^2 + z_1^2 - c^2t_1^2 = 0 \tag{ii}$$

Now the two following postulates are fundamental to the special theory of relativity.

I. A description of a physical phenomenon can contain no reference, either explicitly or implicitly, to the absolute speed of the coordinate frame in which the phenomenon is described.

II. The speed of light is independent of the motion of its source, and is the same for any two observers who are travelling in uniform, rectilinear motion.

Suppose that the relative motions of the two coordinate systems are directed along a common x-axis.

The Galilean transformation from one system to the other is given by

$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \\ t_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & -V \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \tag{iii}$$

where V is the velocity of O_1 relative to O.

By the two postulates above each observer must find the same form of the equation of the wavefront is his own coordinate system.

(*a*) By direct substitution of the Galilean transformation (iii) in equation (ii), show that this cannot be the case.

(*b*) Show that substitution of the Lorentz transformation

$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \\ t_1 \end{pmatrix} = \begin{pmatrix} \dfrac{c}{\sqrt{c^2 - V^2}} & 0 & 0 & \dfrac{-cV}{\sqrt{c^2 - V^2}} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \dfrac{-V}{c\sqrt{c^2 - V^2}} & 0 & 0 & \dfrac{c}{\sqrt{c^2 - V^2}} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}$$

in equation (ii), then the form of the equation does remain the same.

(*c*) Find the inverse Lorentz transformation of the x, y, z, t system into x_1, y_1, z_1, t_1.

(*d*) Show that the transformation in (*b*) above could be described by

$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \\ jct_1 \end{pmatrix} = \begin{pmatrix} \dfrac{c}{\sqrt{c^2 - V^2}} & 0 & 0 & \dfrac{jV}{\sqrt{c^2 - V^2}} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \dfrac{-jV}{\sqrt{c^2 - V^2}} & 0 & 0 & \dfrac{c}{\sqrt{c^2 - V^2}} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ jct \end{pmatrix}$$

where the matrix of the transformation is orthogonal, and $j^2 = -1$.

Hence prove that

$$(x_1 \quad y_1 \quad z_1 \quad jct_1)\begin{pmatrix} x_1 \\ y_1 \\ z_1 \\ jct_1 \end{pmatrix} = (x \quad y \quad z \quad jct)\begin{pmatrix} x \\ y \\ z \\ jct \end{pmatrix}$$

11. (i) If $p, q \in R$, then let the function

$$f(x) = pe^{ax} + qe^{bx}$$

where $a \neq 0$, $b \neq 0$, be represented by $\begin{pmatrix} p \\ q \end{pmatrix}$.

(i) Show that $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}\begin{pmatrix} p \\ q \end{pmatrix}$ represents $f'(x)$.

(ii) Show that $\int f(x)\,dx$ can be represented by

$$\begin{pmatrix} \frac{1}{a} & 0 \\ 0 & \frac{1}{b} \end{pmatrix}\begin{pmatrix} p \\ q \end{pmatrix} + c$$

(iii) What relationship exists between

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \frac{1}{a} & 0 \\ 0 & \frac{1}{b} \end{pmatrix}?$$

(iv) Find $g''(x)$ if $g(x) = 3e^{x} + 2e^{-x}$.
(v) Find $f^{n}(x)$.

12. If $p, q \in R$, then let the function

$$f(x) = pe^{ax}\cos bx + qe^{ax}\sin bx$$

be represented by

$$\begin{pmatrix} p \\ q \end{pmatrix}$$

(i) Show that $f'(x)$ can be represented by

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}\begin{pmatrix} p \\ q \end{pmatrix}$$

(ii) By considering the inverse of the matrix

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

interpret the expression

$$\frac{1}{a^2 + b^2}\begin{pmatrix} a & -b \\ b & a \end{pmatrix}\begin{pmatrix} r \\ s \end{pmatrix}$$

and prove your hypothesis.

(iii) Use your answer to (i) to find

$$\frac{d}{dx}\,[e^{2x}\cos 3x].$$

(iv) Use your answer to (ii) to find

$$\int e^{3x}\sin 2x\,dx.$$

5
Powers of Matrices

5.1 In nearly every situation involving matrices, one problem seems common—that of finding the powers of the matrices. A number of different methods can be used, and we illustrate some of them. For simplicity we shall begin by considering 2 × 2 matrices and extend our ideas as necessary.

Eigenvalues and eigenvectors

5.2 We have seen in the previous chapter (p. 84) that under a matrix transformation the direction of certain column vectors is unchanged; these are called the eigenvectors (or latent vectors or characteristic vectors). We consider the problem algebraically here.

Example

Are there any column vectors whose direction is unchanged under the transformation represented by the matrix

$$\begin{pmatrix} 4 & 2 \\ -3 & -1 \end{pmatrix}?$$

We need a column vector $\begin{pmatrix} x \\ y \end{pmatrix}$ such that

$$\begin{pmatrix} 4 & 2 \\ -3 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = k\begin{pmatrix} x \\ y \end{pmatrix}$$

$$\Leftrightarrow \qquad \left.\begin{aligned} 4x + 2y &= kx \\ -3x - y &= ky \end{aligned}\right\}$$

$$\Leftrightarrow \qquad \left.\begin{aligned} (4 - k)x &= -2y \\ -3x &= (k + 1)y \end{aligned}\right\}$$

These equations are consistent if and only if

$$\frac{4-k}{-3} = \frac{-2}{k+1}$$

$$\begin{aligned} &\Leftrightarrow \quad (4-k)(k+1) = 6 \\ &\Leftrightarrow \quad k^2 - 3k + 2 = 0 \\ &\Leftrightarrow \quad (k-2)(k-1) = 0 \\ &\Leftrightarrow \quad k = 2 \quad \text{or} \quad k = 1 \end{aligned}$$

(i) $k = 2$ gives, from $(4-k)x = -2y$, the relation $x = -y$.
(ii) $k = 1$ gives $3x = -2y$.

We test a particular column vector such that $x = -y$ to check.

Suppose, for example, we consider $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Then

$$\begin{pmatrix} 4 & 2 \\ -3 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = 2\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

and in general

$$\begin{pmatrix} 4 & 2 \\ -3 & -1 \end{pmatrix}\begin{pmatrix} m \\ -m \end{pmatrix} = 2\begin{pmatrix} m \\ -m \end{pmatrix}$$

where m is any scalar. Hence if we can find one column vector which satisfies our requirements, then any scalar multiple of it will do as well.

For the relation $3x = -2y$ we can try the column vector $\begin{pmatrix} 2 \\ -3 \end{pmatrix}$.

Then $$\begin{pmatrix} 4 & 2 \\ -3 & -1 \end{pmatrix}\begin{pmatrix} 2 \\ -3 \end{pmatrix} = 1\begin{pmatrix} 2 \\ -3 \end{pmatrix}$$

5.3 *In general,* numbers k such that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = k\begin{pmatrix} x \\ y \end{pmatrix}$$

are the *eigenvalues* of the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and the corresponding *eigenvectors* are the values of $\begin{pmatrix} x \\ y \end{pmatrix}$ for which the equation holds. Remember that the eigenvectors are by no means unique. Note also that some matrices have no real eigenvalues. We can consider the problem in general. If

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = k\begin{pmatrix} x \\ y \end{pmatrix}$$

$$\Rightarrow A\begin{pmatrix} x \\ y \end{pmatrix} = kI\begin{pmatrix} x \\ y \end{pmatrix} \quad \left[\text{where } I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right]$$

$$\Leftrightarrow (A - kI)\begin{pmatrix} x \\ y \end{pmatrix} = O$$

$$\Leftrightarrow \left\{\begin{pmatrix} a & b \\ c & d \end{pmatrix} - k\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right\}\begin{pmatrix} x \\ y \end{pmatrix} = O$$

$$\Leftrightarrow \begin{pmatrix} a-k & b \\ c & d-k \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = O$$

$$\Leftrightarrow \left.\begin{aligned} (a-k)x + by &= 0 \\ cx + (d-k)y &= 0 \end{aligned}\right\}$$

These equations are consistent if and only if

$$\frac{a-k}{c} = \frac{b}{d-k}$$

$$\Leftrightarrow (a-k)(d-k) - bc = 0$$

$$\Leftrightarrow \Delta \equiv \begin{vmatrix} a-k & b \\ c & d-k \end{vmatrix} = 0^{*}$$

or rearranging

$$\mathbf{k^2 - (a + d)k + ad - bc = 0}$$

5.4 This is called the *characteristic equation* of the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$. If

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

then

$$A^2 = \begin{pmatrix} a^2 + bc & b(a+d) \\ c(a+d) & bc + d^2 \end{pmatrix}$$

$$(a+d)A = \begin{pmatrix} a^2 + ad & ab + bd \\ ac + cd & ad + d^2 \end{pmatrix}$$

so that

$$\underline{\underline{A^2 - (\mathbf{a} + \mathbf{d})A + (\mathbf{ad} - \mathbf{bc})I = 0}} \qquad \text{(i)}$$

Hence the matrix A itself 'satisfies' its own characteristic equation.

---

* N.B. This is the condition for the two non-homogeneous equations to have a non-trivial solution.

This may be extended to larger square matrices and is known as the *Cayley-Hamilton Theorem.* Note that this is *not* a proof of the theorem, just a demonstration that it is true in the case of 2 × 2 matrices.

5.5 Equation (i) is most useful for finding powers of matrices. Suppose we rewrite the equation as

$$A^2 - eA + fI = O$$

i.e. $A^2 = eA - fI$,

$$\text{then } A^3 - eA^2 + fA = O$$

$$\text{so that } A^3 - e(eA - fI) + fA = O$$

i.e. $A^3 = (e^2 - f)A - efI$.

$$\text{Also } A^4 - eA^3 + fA^2 = O$$

$$\text{so that } A^4 - e(e^2 - f)A - e(-efI) + f(eA - fI) = O$$

i.e. $A^4 = (e^3 - 2ef)A + (f^2 - e^2f)I$, and so on.

This means that, in general, the powers of a matrix may be expressed in terms of the matrix itself together with a multiple of the unit matrix, i.e.

$$A^n = gA + hI \tag{ii}$$

where g and h are scalars which have to be determined (this is in general an easy problem as we shall see).

5.6 Relation (ii) is particularly simple in the case where the eigenvalues of the characteristic equation are equal.

Suppose that the equal roots are $k = \lambda$.

Then the Cayley-Hamilton Theorem (for 2 × 2 matrices) gives

$$A^2 - 2\lambda A + \lambda^2 I = O$$

and
$$A^3 - 2\lambda A^2 + \lambda^2 A = O$$

and
$$A^4 - 2\lambda A^3 + \lambda^2 A^2 = O$$

and so on (multiplying each equation by A). On substitution we have

$$\begin{aligned} A^2 &= 2\lambda A - \lambda^2 I \\ A^3 &= 3\lambda^2 A - 2\lambda^3 I \\ A^4 &= 4\lambda^3 A - 3\lambda^4 I \\ &\cdots\cdots\cdots \end{aligned}$$

and in general

$$A^n = n\lambda^{n-1}A - (n-1)\lambda^n I \qquad \text{for } n \geqslant 2$$

which may be proved easily by induction.

5.7 *Note* that the Cayley-Hamilton Theorem also gives us another method for finding the inverse of a matrix. For suppose

$$A^2 - eA + fI = O$$

Then multiplying through by A^{-1} we have

$$A - eI + fA^{-1} = O$$

so that

$$fA^{-1} = eI - A$$

$$\underline{\underline{A^{-1} = \frac{1}{f}\{eI - A\}}}$$

5.8 Eigenvalues and eigenvectors are used to find the powers of a matrix in the following way. Continuing with our initial example where

$$A = \begin{pmatrix} 4 & 2 \\ -3 & -1 \end{pmatrix}$$

and remembering that the eigenvalues were $k = 2$, with corresponding eigenvector $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$, and $k = 1$, with corresponding eigenvector $\begin{pmatrix} 2 \\ -3 \end{pmatrix}$ we now form a new matrix

$$U = \begin{pmatrix} 1 & 2 \\ -1 & -3 \end{pmatrix}$$

with the eigenvectors as its columns.

It is an easy matter to show now that

$$\begin{pmatrix} 4 & 2 \\ -3 & -1 \end{pmatrix}\underset{\uparrow\qquad\uparrow}{\begin{pmatrix} 1 & 2 \\ -1 & -3 \end{pmatrix}} = \begin{pmatrix} 1 & 2 \\ -1 & -3 \end{pmatrix}\underset{\qquad\uparrow}{\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}}$$

This eigenvector corresponds to this eigenvalue

or in general

$$AU = UD$$

where D is a diagonal matrix which has the eigenvalues in its leading diagonal and zeros elsewhere. Now if

$$AU = UD$$

then

$$A = UDU^{-1}$$

and

$$\begin{aligned} A^2 &= UDU^{-1} . UDU^{-1} \\ &= UD^2U^{-1} \\ A^3 &= UD^2U^{-1} . UDU^{-1} \\ &= UD^3U^{-1} \end{aligned}$$

and so on, giving

$$A^n = UD^nU^{-1}$$

Now D^n is easy to evaluate since

$$D = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, \quad D^2 = \begin{pmatrix} 2^2 & 0 \\ 0 & 1^2 \end{pmatrix}, \quad D^3 = \begin{pmatrix} 2^3 & 0 \\ 0 & 1^3 \end{pmatrix}$$

$$D^4 = \begin{pmatrix} 2^4 & 0 \\ 0 & 1^4 \end{pmatrix}, \text{ and so } D^n = \begin{pmatrix} 2^n & 0 \\ 0 & 1^n \end{pmatrix}$$

Hence, for example,

$$\begin{aligned} A^6 &= UD^6U^{-1} \\ &= \begin{pmatrix} 1 & 2 \\ -1 & -3 \end{pmatrix}\begin{pmatrix} 2^6 & 0 \\ 0 & 1^6 \end{pmatrix}\begin{pmatrix} 3 & 2 \\ -1 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 \\ -1 & -3 \end{pmatrix}\begin{pmatrix} 64 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 3 & 2 \\ -1 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 64 & 2 \\ -64 & -3 \end{pmatrix}\begin{pmatrix} 3 & 2 \\ -1 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 190 & 126 \\ -189 & -125 \end{pmatrix} \end{aligned}$$

5.9 This method is usually the easiest one to apply; however there are problems. In the above example U was a non-singular matrix and hence U^{-1} exists, but this is not always the case. If the eigenvalues are equal then the eigenvectors are equal or one eigenvector will be a scalar multiple of the other, e.g.

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} \text{ and } \begin{pmatrix} 2 \\ 4 \end{pmatrix} \left\{ = 2\begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\}$$

This leads to

$$U = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$$

and U is singular, so that U^{-1} does not exist. In this case we can use a variation on our original ideas, using the *Jordan Canonical Form.*

5.10 *Example.* To find the nth power of the matrix

$$B = \begin{pmatrix} 2 & 1 \\ -1 & 4 \end{pmatrix}$$

The characteristic equation is

$$k^2 - 6k + 9 = 0$$

$$\Leftrightarrow k = 3, \quad \text{or} \quad k = 3$$

The idea then is to write

$$B = V\Lambda V^{-1} \qquad (\text{or } BV = V\Lambda)$$

where $$\Lambda = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$$

i.e. a matrix containing the eigenvalues as the leading diagonal with a 1 in the top right hand entry.

Then suppose V is the matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

and we have

$$\begin{pmatrix} 2 & 1 \\ -1 & 4 \end{pmatrix} = \frac{1}{\Delta}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

where $$\Delta = ad - bc$$

so

$$\begin{pmatrix} 2 & 1 \\ -1 & 4 \end{pmatrix} = \frac{1}{\Delta}\begin{pmatrix} 3a & a + 3b \\ 3c & c + 3d \end{pmatrix}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

$$= \frac{1}{\Delta}\begin{pmatrix} 3ad - ac - 3bc & a^2 \\ -c^2 & -3bc + ac + 3ad \end{pmatrix}$$

By comparing entries we have

$$\begin{aligned} 2\Delta &= 3ad + ac - 3bc & \qquad \Delta &= a^2 \\ 4\Delta &= -3bc + ac + 3ad & -\Delta &= -c^2 \end{aligned}$$

Effectively we have three equations,

$$\begin{aligned} \Delta &= ac \\ \Delta &= a^2 \\ -\Delta &= -c^2 \end{aligned}$$

Since V is by no means unique we chose the easiest values for Δ, a and c. Take $\Delta = 1$, then $a = 1$, $c = 1$ are possibilities.

Since $\Delta = 1$, $ad - bc = 1$ so that, if $a = 1$, $c = 1$ then $d - b = 1$ and we could allow for example $d = 3$, $b = 2$.

In this case

$$V = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} \quad \text{and} \quad V^{-1} = \begin{pmatrix} 3 & -2 \\ -1 & 1 \end{pmatrix}$$

To check

$$\begin{aligned} V\Lambda V^{-1} &= \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} 3 & -2 \\ -1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 3 & 7 \\ 3 & 10 \end{pmatrix}\begin{pmatrix} 3 & -2 \\ -1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 1 \\ -1 & 4 \end{pmatrix} \\ &= B. \end{aligned}$$

As previously

$$B^n = V\Lambda^n V^{-1}$$

and our problem now is to find Λ^n

$$\Lambda = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}, \qquad \Lambda^2 = \begin{pmatrix} 3^2 & 6 \\ 0 & 3^2 \end{pmatrix}$$

$$\Lambda^3 = \begin{pmatrix} 3^3 & 27 \\ 0 & 3^3 \end{pmatrix}, \qquad \Lambda^4 = \begin{pmatrix} 3^4 & 108 \\ 0 & 3^4 \end{pmatrix}$$

and so on. The pattern is much easier to recognize in the following form.

$$\Lambda = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}, \qquad \Lambda^2 = \begin{pmatrix} 3^2 & 2.3 \\ 0 & 3^2 \end{pmatrix}$$

$$\Lambda^3 = \begin{pmatrix} 3^3 & 3.3^2 \\ 0 & 3^3 \end{pmatrix} \qquad \Lambda^4 = \begin{pmatrix} 3^4 & 4.3^3 \\ 0 & 3^4 \end{pmatrix}$$

$$\Lambda^5 = \begin{pmatrix} 3^5 & 5.3^4 \\ 0 & 3^4 \end{pmatrix} \cdots$$

In general

$$\Lambda^n = \begin{pmatrix} 3^n & n3^{n-1} \\ 0 & 3^n \end{pmatrix}$$

This may be proved easily using induction. We are now able to find the powers of any 2 × 2 matrix provided that it has real eigenvalues.

Use of Cayley-Hamilton theorem in calculating powers

5.11 Suppose that the matrix A, 'satisfying' its characteristic equation

$$A^2 + lA + mI = O$$

can also be expressed in the form

$$A = UDU^{-1}$$

where U and D are non-singular matrices as explained. By substitution in the characteristic equation for A we have

$$UDU^{-1} . UDU^{-1} + l . UDU^{-1} + mI = O$$

or

$$UD^2 . U^{-1} + l . UDU^{-1} + mI = O$$

Multiply through this equation in front by U^{-1} and behind by U,

$$U^{-1}UD^2U^{-1}U + l . U^{-1}UDU^{-1}U + mU^{-1}IU = O$$

i.e.

$$D^2 \quad + \quad lD \quad + \quad mI \quad = O$$

Hence the diagonal matrix D also 'satisfies' the characteristic equation of the matrix A. This enables us to calculate powers easily and quickly.

Example. To find $\begin{pmatrix} 4 & -3 \\ 1 & 0 \end{pmatrix}^n$. Call $\begin{pmatrix} 4 & -3 \\ 1 & 0 \end{pmatrix}$ the matrix C.

The eigenvalues are $k = 3$ and $k = 1$.
Now using the Cayley-Hamilton theorem,

$$C^n = pC + qI$$

and similarly $D^n = pD + qI$ (as above).

$$D^n = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}^n = p \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} + q \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\Rightarrow \quad \left.\begin{aligned} 3^n &= 3p + q \\ 1^n &= p + q \end{aligned}\right\} \text{ by comparing entries,}$$

so that

$$p = \frac{3^n - 1^n}{2}, \quad q = \frac{3 - 3^n}{2}$$

Hence

$$C^n = \frac{3^n - 1}{2} \begin{pmatrix} 4 & -3 \\ 1 & 0 \end{pmatrix} + \frac{3 - 3^n}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and the problem is solved.

5.12 An easier method for finding the powers of matrices, which works in the cases both when the eigenvalues are different and when the eigenvalues are equal is as follows.

Theorem

$$\begin{pmatrix} p & q \\ r & s \end{pmatrix}^n = k_n \begin{pmatrix} p - \lambda_1 & q \\ r & s - \lambda_2 \end{pmatrix} + \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix}$$

where $k_n = \lambda_1^{n-1} + \lambda_1^{n-2}\lambda_2 + \lambda_1^{n-3}\lambda_2^2 + \ldots + \lambda_2^{n-1}$

This may be proved by induction.

Suppose the formula is correct then

$$\begin{pmatrix} p & q \\ r & s \end{pmatrix}^{n+1} = k_n \begin{pmatrix} p & q \\ r & s \end{pmatrix} \begin{pmatrix} p - \lambda_1 & q \\ r & s - \lambda_2 \end{pmatrix} + \begin{pmatrix} p & q \\ r & s \end{pmatrix} \begin{pmatrix} \lambda_1{}^n & 0 \\ 0 & \lambda_2{}^n \end{pmatrix}$$

but $$\begin{pmatrix} p & q \\ r & s \end{pmatrix} \begin{pmatrix} p - \lambda_1 & q \\ r & s - \lambda_2 \end{pmatrix} = \begin{pmatrix} p - \lambda_1 & q \\ r & s - \lambda_2 \end{pmatrix} \begin{pmatrix} \lambda_2 & 0 \\ 0 & \lambda_1 \end{pmatrix}$$

This may be proved since λ_1, λ_2 are the roots of the characteristic equation $\lambda^2 - (p + s)\lambda + ps - qr = 0$.

Hence

$$\begin{pmatrix} p & q \\ r & s \end{pmatrix}^{n+1} = k_n \begin{pmatrix} p - \lambda_1 & q \\ r & s - \lambda_2 \end{pmatrix} \begin{pmatrix} \lambda_2 & 0 \\ 0 & \lambda_1 \end{pmatrix} + \begin{pmatrix} p & q \\ r & s \end{pmatrix} \begin{pmatrix} \lambda_1{}^n & 0 \\ 0 & \lambda_2{}^n \end{pmatrix}$$

$$= \begin{pmatrix} p - \lambda_1 & q \\ r & s - \lambda_2 \end{pmatrix} \left\{ k_n \begin{pmatrix} \lambda_2 & 0 \\ 0 & \lambda_1 \end{pmatrix} + \begin{pmatrix} \lambda_1{}^n & 0 \\ 0 & \lambda_2{}^n \end{pmatrix} \right\} + \begin{pmatrix} \lambda_1{}^{n+1} & 0 \\ 0 & \lambda_2{}^{n+1} \end{pmatrix}$$

But $$k_n \begin{pmatrix} \lambda_2 & 0 \\ 0 & \lambda_1 \end{pmatrix} + \begin{pmatrix} \lambda_1{}^n & 0 \\ 0 & \lambda_2{}^n \end{pmatrix} = k_{n+1} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

{since $k_{n+1} = \lambda_1{}^n + \lambda_1{}^{n-1}\lambda_2 + \ldots \lambda_2{}^n$}

so $$\begin{pmatrix} p & q \\ r & s \end{pmatrix}^{n+1} = k_{n+1} \begin{pmatrix} p - \lambda_1 & q \\ r & s - \lambda_2 \end{pmatrix} + \begin{pmatrix} \lambda_1{}^{n+1} & 0 \\ 0 & \lambda_2{}^{n+1} \end{pmatrix}$$

This is clearly true for $n = 1$ and hence $n = 2, 3 \ldots$

Exercise 5.1

Questions 1–4 involve matrices with distinct eigenvalues, Questions 5–8 involve matrices with equal eigenvalues.

1. If M is the matrix $\begin{pmatrix} 5 & 2 \\ 3 & 4 \end{pmatrix}$, find the values of

   (i) M^5, (ii) M^n.

2. Find the value of P^n where

$$P = \begin{pmatrix} \frac{1}{5} & \frac{4}{5} \\ \frac{2}{5} & \frac{3}{5} \end{pmatrix}$$

and find the limit of P^n as $n \to \infty$.

3. Find Q^6 and Q^n where

$$Q = \begin{pmatrix} -\frac{1}{6} & 1 \\ -\frac{1}{3} & 1 \end{pmatrix}$$

4. Find R^7 and R^n where

$$R = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}$$

5. Find S^5 and S^n where

$$S = \begin{pmatrix} 6 & -1 \\ 1 & 4 \end{pmatrix}$$

6. Find T^4 and T^n where

$$T = \begin{pmatrix} -6 & -2 \\ 2 & -2 \end{pmatrix}$$

7. Find W^6 and W^n where

$$W = \begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix}$$

8. Find X^7 and X^n where

$$X = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}$$

9. If $A = \begin{pmatrix} -2 & -9 \\ 1 & 4 \end{pmatrix}$, prove by induction that

$A^n = \begin{pmatrix} 1 - 3n & -9n \\ n & 1 + 3n \end{pmatrix}$ where n is a positive integer.

Write down the matrix which is $\lim\limits_{n \to \infty} \dfrac{1}{n} A^n$. (*MME*)

10. Prove that, when n is a positive integer

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}^n = \begin{pmatrix} \cos n\theta & -\sin n\theta \\ \sin n\theta & \cos n\theta \end{pmatrix}$$

11. Prove that

$$\begin{pmatrix} a & b \\ ka & kb \end{pmatrix}^n = (a + kb)^{n-1} \, . \begin{pmatrix} a & b \\ ka & kb \end{pmatrix}$$

12. Show that if

$$A = \begin{pmatrix} \cosh a & b \sinh a \\ \frac{1}{b} \sinh a & \cosh a \end{pmatrix}$$

then $$A^n = \begin{pmatrix} \cosh na & b \sinh na \\ \frac{1}{b} \sinh na & \cosh na \end{pmatrix}$$

An iterative process for finding $\sqrt{2}$ using matrices

5.13 Consider $\begin{pmatrix}1 & 2\\1 & 1\end{pmatrix}$ and guess the square root of 2 as 4 say, i.e. our first guess $g_0 = 4$, which we write $\begin{pmatrix}4\\1\end{pmatrix}$.

$$\text{Then} \quad \begin{pmatrix}1 & 2\\1 & 1\end{pmatrix}\begin{pmatrix}4\\1\end{pmatrix} = \begin{pmatrix}6\\5\end{pmatrix} \quad \text{and} \quad g_1 = \frac{6}{5} = 1{\cdot}2$$

$$\text{Now} \quad \begin{pmatrix}1 & 2\\1 & 1\end{pmatrix}\begin{pmatrix}6\\5\end{pmatrix} = \begin{pmatrix}16\\11\end{pmatrix} \quad \text{and} \quad g_2 = \frac{16}{11} \approx 1{\cdot}45$$

$$\begin{pmatrix}1 & 2\\1 & 1\end{pmatrix}\begin{pmatrix}16\\11\end{pmatrix} = \begin{pmatrix}38\\27\end{pmatrix} \quad \text{and} \quad g_3 = \frac{38}{27} \approx 1{\cdot}407$$

$$\begin{pmatrix}1 & 2\\1 & 1\end{pmatrix}\begin{pmatrix}38\\27\end{pmatrix} = \begin{pmatrix}92\\65\end{pmatrix} \quad \text{and} \quad g_4 = \frac{92}{65} \approx 1{\cdot}415$$

It is an interesting exercise to prove that the sequence g_0, g_1, g_2, g_3, g_4 . . . is converging to $\sqrt{2}$.*

This is an interesting side of the problem of finding eigenvalues by iterative procedures. In the above case one of the eigenvectors will be found to be $\begin{pmatrix}\sqrt{2}\\1\end{pmatrix}$.

To find the eigenvalues and eigenvectors using iteration

5.14 Consider again $\begin{pmatrix}1 & 2\\1 & 1\end{pmatrix}$ with our first guess as $\begin{pmatrix}4\\1\end{pmatrix}$. Then

$$\begin{pmatrix}1 & 2\\1 & 1\end{pmatrix}\begin{pmatrix}4\\1\end{pmatrix} = \begin{pmatrix}6\\5\end{pmatrix} = 5\begin{pmatrix}1{\cdot}2\\1\end{pmatrix}$$

Instead of applying $\begin{pmatrix}1 & 2\\1 & 1\end{pmatrix}$ to $\begin{pmatrix}6\\5\end{pmatrix}$ as before, we use $\begin{pmatrix}\frac{6}{5}\\1\end{pmatrix}$,

$$\begin{pmatrix}1 & 2\\1 & 1\end{pmatrix}\begin{pmatrix}\frac{6}{5}\\1\end{pmatrix} = \begin{pmatrix}3{\cdot}2\\2{\cdot}2\end{pmatrix} = 2{\cdot}2\begin{pmatrix}\frac{16}{11}\\1\end{pmatrix} \approx 2{\cdot}2\begin{pmatrix}1{\cdot}45\\1\end{pmatrix}$$

$$\begin{pmatrix}1 & 2\\1 & 1\end{pmatrix}\begin{pmatrix}\frac{16}{11}\\1\end{pmatrix} \approx \begin{pmatrix}3{\cdot}45\\2{\cdot}45\end{pmatrix} \approx 2{\cdot}45\begin{pmatrix}\frac{3{\cdot}45}{2{\cdot}45}\\1\end{pmatrix} \approx 2{\cdot}45\begin{pmatrix}1{\cdot}407\\1\end{pmatrix}$$

$$\begin{pmatrix}1 & 2\\1 & 1\end{pmatrix}\begin{pmatrix}1{\cdot}407\\1\end{pmatrix} \approx \begin{pmatrix}3{\cdot}407\\2{\cdot}407\end{pmatrix} \approx 2{\cdot}407\begin{pmatrix}1{\cdot}415\\1\end{pmatrix}$$

and so on.

---

* A solution is found in *Mathematical Prima No.* 19 (from St. Dunstan's College Mathematical Society.)

The sequence 5, 2·2, 2·45, 2·407 converges to give the eigenvalue whilst the other sequence 4, 1·2, 1·45, 1·407, 1·415 converges to give the corresponding eigenvector.

5.15 *Justification.* Consider a guess $\mathbf{g}_1$, say, then

$$\mathbf{g}_1 = c_1\mathbf{x}_1 + c_2\mathbf{x}_2$$

where $\mathbf{x}_1$ and $\mathbf{x}_2$ are the two eigenvectors corresponding to λ_1 and λ_2 (the eigenvalues) and c_1, c_2 are scalars. Now

$$A\mathbf{g}_1 = c_1A\mathbf{x}_1 + c_2A\mathbf{x}_2 \qquad \text{and} \qquad \begin{aligned} A\mathbf{x}_1 &= \lambda_1\mathbf{x}_1 \\ A\mathbf{x}_2 &= \lambda_2\mathbf{x}_2 \end{aligned}$$

so

$$\begin{aligned} A\mathbf{g}_1 &= c_1\lambda_1\mathbf{x}_1 + c_2\lambda_2\mathbf{x}_2 \\ &= m\left(\frac{\lambda_1}{m}c_1\mathbf{x}_1 + \frac{\lambda_2}{m}c_2\mathbf{x}_2\right) \\ &= m\mathbf{g}_2 \end{aligned}$$

where m was chosen as shown in the example above.

If $\lambda_2 > \lambda_1$ the contribution from $\mathbf{x}_1$ to the guess $\mathbf{g}_2$ is less important than it was in $\mathbf{g}_1$ and eventually $\mathbf{x}_2$ is dominant (provided that $c_2 \neq 0$).

To find the other (smaller) eigenvalue

5.16 Remember that $\lambda_1 + \lambda_2 = a + d$ (the *trace* of the matrix—the sum of the elements in the leading diagonal) or since

$$A\mathbf{x} = \lambda\mathbf{x}$$

$$\Rightarrow A^{-1}\mathbf{x} = \frac{1}{\lambda}\mathbf{x}$$

use A^{-1} to give the larger value of $\frac{1}{\lambda}$ or the smaller value of λ.

5.17 This idea leads to an **iterative process** for solving quadratic equations.

Example

Solve the equation

$$\lambda^2 - 5\lambda + 2 = 0$$

Consider this equation as the characteristic equation of some matrix A say. Then if

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

we have $a + d = 5$, $ad - bc = 2$.

Let $a = 5$, $d = 0$, $b = 2$, $c = -1$, this is one set of possible values for a, b, c and d. Then

$$A = \begin{pmatrix} 5 & 2 \\ -1 & 0 \end{pmatrix}$$

We then proceed as before to find the eigenvalues (the roots of our equation) and the eigenvectors. Consider as a first guess

$$\mathbf{g}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Then $$\begin{pmatrix} 5 & 2 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 7 \\ -1 \end{pmatrix} = -1\begin{pmatrix} -7 \\ 1 \end{pmatrix}$$

try $$\mathbf{g}_2 = \begin{pmatrix} -7 \\ -1 \end{pmatrix}$$

$$\begin{pmatrix} 5 & 2 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} -7 \\ 1 \end{pmatrix} = \begin{pmatrix} -33 \\ 7 \end{pmatrix} \approx 7\begin{pmatrix} -4{\cdot}7 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 5 & 2 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} -4{\cdot}7 \\ 1 \end{pmatrix} = \begin{pmatrix} -21{\cdot}5 \\ 4{\cdot}7 \end{pmatrix} \approx 4{\cdot}7\begin{pmatrix} -4{\cdot}57 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 5 & 2 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} -4{\cdot}57 \\ 1 \end{pmatrix} = \begin{pmatrix} -20{\cdot}85 \\ 4{\cdot}57 \end{pmatrix} \approx 4{\cdot}57\begin{pmatrix} -4{\cdot}56 \\ 1 \end{pmatrix}$$

It is now clear that the largest eigenvalue lies between 4·56 and 4·57. The smaller one can be found easily since the sum of the roots is 5.

Check

Using the formula for the solution of quadratic equations

$$\lambda = \frac{5 \pm \sqrt{25 - 8}}{2}$$

$$= \frac{5 \pm \sqrt{17}}{2}$$

The largest value of λ is thus

$$\frac{5 + \sqrt{17}}{2} \approx \frac{5 + 4{\cdot}123}{2}$$

$$\approx \underline{\underline{4{\cdot}561(5)}}$$

Further consideration of the problem of finding roots using matrices.*

5.18 The difficulty of the process shown is that the series does not converge to give $\sqrt{2}$ very quickly. However it is worth considering for yourself what happens if the first guess is a wild one (4000 say). This process does not give the negative square root.

Consider again

$$\begin{pmatrix}1&2\\1&1\end{pmatrix}^2=\begin{pmatrix}3&4\\2&3\end{pmatrix}\qquad \text{Mean of } \tfrac{3}{2} \text{ and } \tfrac{4}{3} = 1{\cdot}415$$

$$\begin{pmatrix}3&4\\2&3\end{pmatrix}^2=\begin{pmatrix}17&24\\12&17\end{pmatrix}\qquad \text{Mean of } \tfrac{17}{12} \text{ and } \tfrac{24}{17} = 1{\cdot}4142$$

Consider also

$$\begin{pmatrix}5&2\\1&5\end{pmatrix}^2=\begin{pmatrix}27&20\\10&27\end{pmatrix}\qquad \text{Mean of } \tfrac{27}{10} \text{ and } \tfrac{20}{27} = 1{\cdot}72$$

$$\begin{pmatrix}27&20\\10&27\end{pmatrix}^2=\begin{pmatrix}929&1080\\540&929\end{pmatrix}\qquad \text{Mean of } \tfrac{929}{540} \text{ and } \tfrac{1080}{929} = 1{\cdot}442$$

This process should be compared with the usual iterative process for finding square roots, i.e. $x_{r+1} = \frac{1}{2}\left(x_r + \frac{N}{x_r}\right)$.

How could we find $\sqrt{-2}$?

Consider

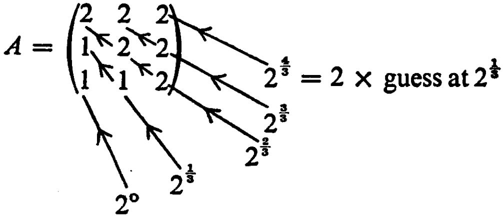
$$\begin{pmatrix}j&-2\\1&j\end{pmatrix}^2=\begin{pmatrix}-3&-4j\\2j&-3\end{pmatrix}$$

$$\begin{pmatrix}-3&-4j\\2j&-3\end{pmatrix}^2=\begin{pmatrix}17&-24j\\-12j&17\end{pmatrix}$$

The ratios clearly tend to $\sqrt{2}\,.\,j$.

Examples of other uses

1. To find $2^{\frac{1}{3}}$


$A = \begin{pmatrix}2&2&2\\1&2&2\\1&1&2\end{pmatrix}$
$2^{\frac{4}{3}} = 2 \times$ guess at $2^{\frac{1}{3}}$
$2^{\frac{3}{3}}$
$2^{\frac{2}{3}}$
$2^{\frac{1}{3}}$
2^0

---

* We are indebted to Alastair McDougall for these ideas.

$$A^2 = \begin{pmatrix} 8 & 10 & 12 \\ 6 & 8 & 10 \\ 5 & 6 & 8 \end{pmatrix}$$ We allow no leeway in 2^0 or $2^{\frac{3}{3}}$

New guess $\frac{1}{3}[\frac{6}{5} + \frac{8}{6} + \frac{10}{8}]$

1·4
1·33
1·25

$3\overline{)3{\cdot}98}$

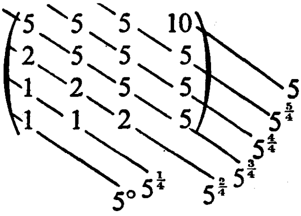
The mean is 1·33

$$A^4 = \begin{pmatrix} 184 & 232 & 292 \\ 146 & 184 & 232 \\ 116 & 146 & 184 \end{pmatrix}$$

Our estimates of $2^{\frac{1}{3}}$, $2^{\frac{2}{3}}$ are now $\frac{146}{116}$, $\frac{184}{116}$ respectively. Three estimates for $2^{\frac{1}{3}}$ can be got from A^4

$$\left.\begin{aligned} \tfrac{146}{116} &= 1{\cdot}2586207 \\ \tfrac{184}{146} &= 1{\cdot}2602740 \\ \tfrac{232}{184} &= 1{\cdot}2608696 \end{aligned}\right\}$$ The Arithmetic mean gives $2^{\frac{1}{3}} = 1{\cdot}2599214$ (An error of 3 in the last digit)

2. To find $5^{\frac{1}{4}}$



$$A^2 = \begin{pmatrix} 50 & 70 & 95 & 150 \\ 30 & 50 & 70 & 95 \\ 19 & 30 & 52 & 20 \\ 14 & 19 & 30 & 50 \end{pmatrix}$$

Our estimates are

$$\left.\begin{aligned} \tfrac{19}{14} &= 1{\cdot}357 \\ \tfrac{30}{19} &= 1{\cdot}579 \\ \tfrac{50}{30} &= 1{\cdot}667 \\ \tfrac{70}{50} &= 1{\cdot}400 \end{aligned}\right\}$$ The Arithmetic mean is 1·501 (An error of 0·006 or 0·4%)

3. To find the complex cube roots of 1 say, it is necessary to 'prime' the process by feeding in complex or imaginary estimates.

We choose j, as our estimate for $1^{\frac{1}{3}}$, $-j$, for $1^{\frac{2}{3}}$,

$$\begin{pmatrix} -j & 1 & j \\ j & -j & 1 \\ 1 & j & -j \end{pmatrix}^4 = \begin{pmatrix} -1+2j & -1-2j & 3 \\ 3 & -1+2j & -1-2j \\ -1-2j & 3 & -1+2j \end{pmatrix}^2$$

$$= \begin{pmatrix} -9-16j & 19 & \dots \\ -9+16j & \dots & \\ 19 & \dots & \end{pmatrix}$$

Taking the mean of $\dfrac{-9+16j}{19}, \dfrac{-9-16j}{-9+16j}, \dfrac{19}{-9-16j}$

we get $-0{\cdot}5001 + 0{\cdot}8663j$ (giving an error of about $0{\cdot}03\%$)

Exercise 5.2

Use the iterative processes described to find

1. $\sqrt{5}$ taking (*a*) as your first guess 3, (*b*) as your first guess 100,000. (Find $\sqrt{5}$ to two-decimal places in each case.)

2. Repeat Question 1 for $\sqrt{7}$.

Solve the equations

3. $\lambda^2 - 2\lambda - 5 = 0$.
4. $3\lambda^2 - 7\lambda + 1 = 0$.

5. What happens if the quadratic equation given has complex solutions? Investigate the solution of

$$\lambda^2 - \lambda + 3 = 0$$

Powers of 3×3 matrices and 4×4 matrices

5.19 The methods given for 2 × 2 matrices may easily be extended to deal with larger matrices. In general the ideas are the same although the algebra is often more complicated.

Example

Find P^n where

$$P = \begin{pmatrix} 3 & 2 & 2 \\ 1 & 4 & 1 \\ -2 & -4 & -1 \end{pmatrix}$$

To find the eigenvalues, remember, that the determinant $|P - \lambda I| = 0$, i.e.

$$\begin{vmatrix} 3-\lambda & 2 & 2 \\ 1 & 4-\lambda & 1 \\ -2 & -4 & -1-\lambda \end{vmatrix} = 0$$

Multiplying out and simplifying we have

$$\begin{aligned} \lambda^3 - 6\lambda^2 + 11\lambda - 6 &= 0 \\ \Leftrightarrow (\lambda - 1)(\lambda - 2)(\lambda - 3) &= 0 \\ \Leftrightarrow \lambda = 1, \text{ or } \lambda = 2, &\text{ or } \lambda = 3. \end{aligned}$$

Consider when $\lambda = 1$ and we have an eigenvector $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$. Then

$$\begin{pmatrix} 3 & 2 & 2 \\ 1 & 4 & 1 \\ -2 & -4 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 1 \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$\left.\begin{aligned} 3x + 2y + 2z &= x \\ x + 4y + z &= y \\ -2x - 4y - z &= z \end{aligned}\right\} \Rightarrow \left\{\begin{aligned} x + y + z &= 0 \\ x + 3y + z &= 0 \\ x + 2y + z &= 0 \end{aligned}\right.$$

Hence $y = 0$ and we can take $z = 1$, $x = -1$.

An eigenvector is in this case

$$\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

when $\lambda = 2$. Then

$$\begin{pmatrix} 3 & 2 & 2 \\ 1 & 4 & 1 \\ -2 & -4 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 2 \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$\left.\begin{aligned} 3x + 2y + 2z &= 2x \\ x + 4y + z &= 2y \\ -2x - 4y - 4z &= 2z \end{aligned}\right\} \Rightarrow \left\{\begin{aligned} x + 2y + 2z &= 0 \\ x + 2y + z &= 0 \\ x + 2y + 3z &= 0 \end{aligned}\right.$$

Hence $z = 0$ and we can take $y = 1$, $x = -2$.

This time an eigenvector is

$$\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}$$

when $\lambda = 3$. Then

$$\begin{pmatrix} 3 & 2 & 2 \\ 1 & 4 & 1 \\ -2 & -4 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 3\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$\left.\begin{aligned} 3x + 2y + 2z &= 3x \\ x + 4y + z &= 3y \\ -2x - 4y - z &= 3z \end{aligned}\right\} \Rightarrow \left.\begin{aligned} y + z &= 0 \\ x + y + z &= 0 \\ \tfrac{1}{2}x + y + z &= 0 \end{aligned}\right\}$$

Hence $x = 0$ and we can take $z = 1, y = -1$.

An eigenvector is

$$\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$$

To find P^n we now use the idea, as before, that we can write $PU = UD$ where U and D are defined as previously, so that

$$\begin{aligned} P^n &= UD^nU^{-1} \\ &= \begin{pmatrix} -1 & -2 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1^n & 0 & 0 \\ 0 & 2^n & 0 \\ 0 & 0 & 3^n \end{pmatrix} . U^{-1} \end{aligned}$$

where the columns of U are the eigenvectors corresponding in order to the entries in D.

We now need to find U^{-1} if it exists. This can be done using one of the methods described in Chapter 2.

It is given by

$$\tfrac{1}{3}\begin{pmatrix} 1 & 2 & 2 \\ -1 & 1 & 1 \\ -1 & -2 & 1 \end{pmatrix}$$

Hence

$$\begin{aligned} P^n &= \tfrac{1}{3}\begin{pmatrix} -1 & -2 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1^n & 0 & 0 \\ 0 & 2^n & 0 \\ 0 & 0 & 3^n \end{pmatrix}\begin{pmatrix} 1 & 2 & 2 \\ -1 & 1 & 1 \\ -1 & -2 & 1 \end{pmatrix} \\ &= \tfrac{1}{3}\begin{pmatrix} -1 & -2\cdot 2^n & 0 \\ 0 & 2^n & -3^n \\ 1 & 0 & 3^n \end{pmatrix}\begin{pmatrix} 1 & 2 & 2 \\ -1 & 1 & 1 \\ -1 & -2 & 1 \end{pmatrix} \\ &= \tfrac{1}{3}\begin{pmatrix} -1 + 2^{n+1} & -2(2^n + 1) & -2(2^n + 1) \\ -2^n - 3^n & 2^n + 2.3^n & 2^n - 3^n \\ 1 - 3^n & 2(1 - 3^n) & 2 + 3^n \end{pmatrix} \end{aligned}$$

5.20 It is possible, in general, to use an iterative process for finding the eigenvalues and eigenvectors of a 3 × 3 matrix.

Example

To find the eigenvalues of

$$\begin{pmatrix} 3 & 2 & 2 \\ 1 & 4 & 1 \\ -2 & -4 & -1 \end{pmatrix}$$

Guess an eigenvector $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ say, then

$$\begin{pmatrix} 3 & 2 & 2 \\ 1 & 4 & 1 \\ -2 & -4 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 7 \\ 6 \\ -7 \end{pmatrix} \approx -7\begin{pmatrix} -1 \\ -0{\cdot}9 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 3 & 2 & 2 \\ 1 & 4 & 1 \\ -2 & -4 & -1 \end{pmatrix}\begin{pmatrix} -1 \\ -0{\cdot}9 \\ 1 \end{pmatrix} = \begin{pmatrix} -2{\cdot}8 \\ -3{\cdot}6 \\ 4{\cdot}6 \end{pmatrix} \approx 4{\cdot}6\begin{pmatrix} -0{\cdot}6 \\ -0{\cdot}8 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 3 & 2 & 2 \\ 1 & 4 & 1 \\ -2 & -4 & -1 \end{pmatrix}\begin{pmatrix} -0{\cdot}6 \\ -0{\cdot}8 \\ 1 \end{pmatrix} = \begin{pmatrix} -1{\cdot}2 \\ -2{\cdot}8 \\ 3{\cdot}4 \end{pmatrix} \approx 3{\cdot}4\begin{pmatrix} -0{\cdot}4 \\ -0{\cdot}8 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 3 & 2 & 2 \\ 1 & 4 & 1 \\ -2 & -4 & -1 \end{pmatrix}\begin{pmatrix} -0{\cdot}4 \\ -0{\cdot}8 \\ 1 \end{pmatrix} = \begin{pmatrix} -0{\cdot}8 \\ -2{\cdot}6 \\ 3 \end{pmatrix} \approx 3\begin{pmatrix} -0{\cdot}3 \\ -0{\cdot}9 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 3 & 2 & 2 \\ 1 & 4 & 1 \\ -2 & -4 & -1 \end{pmatrix}\begin{pmatrix} -0{\cdot}3 \\ -0{\cdot}9 \\ 1 \end{pmatrix} = \begin{pmatrix} -0{\cdot}7 \\ -2{\cdot}9 \\ 3{\cdot}2 \end{pmatrix} \approx 3{\cdot}2\begin{pmatrix} -0{\cdot}2 \\ -0{\cdot}9 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 3 & 2 & 2 \\ 1 & 4 & 1 \\ -2 & -4 & -1 \end{pmatrix}\begin{pmatrix} -0{\cdot}2 \\ -0{\cdot}9 \\ 1 \end{pmatrix} = \begin{pmatrix} -0{\cdot}4 \\ -2{\cdot}8 \\ 3 \end{pmatrix} \approx 3\begin{pmatrix} -0{\cdot}1 \\ -0{\cdot}9 \\ 1 \end{pmatrix}$$

and finally

$$\begin{pmatrix} 3 & 2 & 2 \\ 1 & 4 & 1 \\ -2 & -4 & -1 \end{pmatrix}\begin{pmatrix} -0{\cdot}1 \\ -0{\cdot}9 \\ 1 \end{pmatrix} = \begin{pmatrix} -0{\cdot}1 \\ -2{\cdot}7 \\ 2{\cdot}8 \end{pmatrix} \approx 2{\cdot}8\begin{pmatrix} 0 \\ -0{\cdot}96 \\ 1 \end{pmatrix}$$

We hope that it is possible to recognize then the eigenvalue (actually 3) and the eigenvector—actually $\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$. This gives the largest eigenvalue. The smallest one can be found using the same process on the inverse matrix and the third using the trace.

5.21 Certain 4 × 4 matrices occur quite frequently when dealing with probability matrices and we consider this special type (an example is given later, see p. 140). These are 4 × 4 matrices which may be partitioned into four 2 × 2 matrices to give, for example

$$Q = \left(\begin{array}{c|c} E & F \\ \hline O & I \end{array}\right)$$

where O is $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ and I is $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Then

$$Q^2 = \begin{pmatrix} E & F \\ O & I \end{pmatrix}\begin{pmatrix} E & F \\ O & I \end{pmatrix}$$

$$= \begin{pmatrix} E^2 & EF + F \\ O & I \end{pmatrix}$$

$$Q^3 = \begin{pmatrix} E^3 & E^2F + EF + F \\ O & I \end{pmatrix}$$

$$Q^4 = \begin{pmatrix} E^4 & E^3F + E^2F + EF + F \\ O & I \end{pmatrix}$$

and so on

$$Q^n = \begin{pmatrix} E^n & E^{n-1}F + E^{n-2}F + \ldots + F \\ O & I \end{pmatrix}$$

$$= \begin{pmatrix} E^n & (E^{n-1} + E^{n-2} + \ldots + I)F \\ O & I \end{pmatrix}$$

$$= \begin{pmatrix} E^n & (I - E^n)(I - E)^{-1}F \\ O & I \end{pmatrix}$$

This follows since the top right hand entry is a geometric progression,

$$\left[\text{cf.} \quad s = \frac{a(1 - r^n)}{1 - r}\right]$$

APPLICATIONS

Apart from a number of topics mentioned in the later sections of this book, one of most important applications—or situations where powers of a matrix is required—is in the realm of *probability.*

5.22 *Example 1.* If a person catches a bus one day there is a probability of $\frac{1}{5}$ that he will catch it the next day. On the other hand there is a probability of $\frac{3}{5}$ that if he misses the bus one day he will also miss it on the second day.

We can display this information as a matrix

$$\begin{array}{ll} & \qquad\quad \text{Second day} \\ & \begin{array}{cc} \textit{Catches bus} & \textit{Misses bus} \end{array} \\ \text{First day} \begin{array}{r} \textit{Catches bus} \\ \textit{Misses bus} \end{array} & \begin{pmatrix} \frac{1}{5} & \frac{4}{5} \\ \frac{2}{5} & \frac{3}{5} \end{pmatrix} = P. \end{array}$$

It can be shown (*Matrices 2* Ch. III) that P^2 will give the probabilities for the 3rd day, e.g.

$$\begin{array}{ll} & \qquad\quad \text{Third day} \\ & \begin{array}{cc} \textit{Catches bus} & \textit{Misses bus} \end{array} \\ \text{First day} \begin{array}{r} C \\ M \end{array} & \begin{pmatrix} \frac{9}{25} & \frac{16}{25} \\ \frac{8}{25} & \frac{17}{25} \end{pmatrix} \end{array}$$

This means, for example, that if the bus is caught on the first day there is a probability of $\frac{9}{25}$ that it will be caught on the third day and so on. We can continue this process:

$$\left[P^2 = \begin{pmatrix} 0{\cdot}36 & 0{\cdot}64 \\ 0{\cdot}32 & 0{\cdot}68 \end{pmatrix}, \; P^3 = \begin{pmatrix} 0{\cdot}328 & 0{\cdot}672 \\ 0{\cdot}336 & 0{\cdot}664 \end{pmatrix}, \; P^4 = \begin{pmatrix} 0{\cdot}3344 & 0{\cdot}6656 \\ 0{\cdot}3328 & 0{\cdot}6672 \end{pmatrix}\right]$$

It is clear that the two rows of our matrix are becoming closer and closer together and the probabilities are depending less and less on the initial probabilities. Consider a long time ahead when the probability of catching the bus is nearly the same whether it was caught or missed on the first day. Let this probability be p.

$$\begin{array}{ll} & \text{A large number of days ahead} \\ & \begin{array}{cc} \textit{Catches bus} & \textit{Misses bus} \end{array} \\ \text{First day} \begin{array}{r} C \\ M \end{array} & \begin{pmatrix} p & 1-p \\ p & 1-p \end{pmatrix} \end{array}$$

As we have chosen a day so far ahead that these probabilities are almost equal, then the probabilities for the day after this must also be approximately the same.

Passing to the limit, when $n \to \infty$,

$$\begin{pmatrix} p & 1-p \\ p & 1-p \end{pmatrix}\begin{pmatrix} \frac{1}{5} & \frac{4}{5} \\ \frac{2}{5} & \frac{3}{5} \end{pmatrix} = \begin{pmatrix} p & 1-p \\ p & 1-p \end{pmatrix}$$

$$\Leftrightarrow \tfrac{1}{5}p + \tfrac{2}{5}(1-p) = p$$

$$\Leftrightarrow p = \tfrac{1}{3}$$

Hence

$$P^n \approx \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}$$

when n is large.

5.25 *Example* 2. Three dogs A, B and C are all fighting for the same bone. When A has it there is a probability of $\frac{1}{2}$ that he will keep it for 5 minutes, and a probability of $\frac{1}{4}$ that he will lose it to B (and similarly for C). The other probabilities are given in the matrix

$$\begin{array}{cc} & \text{Owner of bone at time } t+5 \\ & \begin{array}{ccc} \text{A} & \text{B} & \text{C} \end{array} \\ \begin{array}{c} \text{Owner of bone} \\ \text{at time } t \end{array} \begin{array}{c} \text{A} \\ \text{B} \\ \text{C} \end{array} & \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{3}{4} & 0 \\ \frac{1}{4} & 0 & \frac{3}{4} \end{pmatrix} \end{array}$$

If A has the bone initially, find which dog is likely to have the bone (*a*) after 10 minutes, (*b*) after 20 minutes, (*c*) eventually.

Call the matrix

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{3}{4} & 0 \\ \frac{1}{4} & 0 & \frac{3}{4} \end{pmatrix}$$

The solution requires us to find (i) P^2, (ii) P^4, (iii) P^n.

(i)

$$P^2 = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{3}{4} & 0 \\ \frac{1}{4} & 0 & \frac{3}{4} \end{pmatrix}\begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{3}{4} & 0 \\ \frac{1}{4} & 0 & \frac{3}{4} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{3}{8} & \frac{5}{16} & \frac{5}{16} \\ \frac{5}{16} & \frac{10}{16} & \frac{1}{16} \\ \frac{5}{16} & \frac{1}{16} & \frac{10}{16} \end{pmatrix}$$

Hence if A has the bone initially, the first row of the matrix gives us that the probability of A still having the bone is $\frac{3}{8}$, of B having it $\frac{5}{16}$, of C having it $\frac{5}{16}$—so A is most likely to still have the bone.

(ii) Now

$$P^4 = P^2 . P^2$$

$$= \begin{pmatrix} \frac{3}{8} & \frac{5}{16} & \frac{5}{16} \\ \frac{5}{16} & \frac{10}{16} & \frac{1}{16} \\ \frac{5}{16} & \frac{1}{16} & \frac{10}{16} \end{pmatrix} \begin{pmatrix} \frac{3}{8} & \frac{5}{16} & \frac{5}{16} \\ \frac{5}{16} & \frac{10}{16} & \frac{1}{16} \\ \frac{5}{16} & \frac{1}{16} & \frac{10}{16} \end{pmatrix}$$

$$= \frac{1}{16^2} \begin{pmatrix} 86 & 85 & 85 \\ 85 & 126 & 45 \\ 85 & 45 & 126 \end{pmatrix}$$

or
$$P^4 \approx \begin{pmatrix} 0{\cdot}3359 & 0{\cdot}3321 & 0{\cdot}3321 \\ 0{\cdot}3321 & 0{\cdot}4922 & 0{\cdot}1758 \\ 0{\cdot}3321 & 0{\cdot}1758 & 0{\cdot}4922 \end{pmatrix}$$

It is becoming clear that as we continue (as in the previous example) the entries in our matrix tend to certain limits. Here, as $n \to \infty$,

$$P^n = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} \quad \text{not} \quad \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{2} \end{pmatrix}$$

as it would appear!

Find P^n for yourself, using the idea that

$$P^n = UD^nU^{-1}$$

The eigenvalues are $\frac{1}{4}$, $\frac{3}{4}$ and 1. Hence after 20 minutes (given that A had the bone initially), A is still the most likely dog to have it, and eventually they each have an equal probability of having it (although it appears that A's chances are always slightly the better).

5.23 *Example 3.* Two criminals A and B have just robbed a bank and are awaiting their getaway car. A squad car full of policemen arrives on the scene. The probabilities of the criminals being free for a period of 1 minute are shown in the following matrix,

$$\begin{array}{ll} & \text{FREE AT TIME } t+1 \\ & \begin{array}{cccc} \text{A,B} & \text{A} & \text{B} & \text{nil} \end{array} \\ \text{FREE AT TIME } t \quad \begin{array}{r} \text{A,B} \\ \text{A} \\ \text{B} \\ \text{nil} \end{array} & \begin{pmatrix} \frac{3}{8} & \frac{1}{4} & \frac{1}{4} & \frac{1}{8} \\ \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{2} \\ \frac{1}{4} & 0 & \frac{1}{2} & \frac{1}{4} \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

Make sure that you understand the meaning of the entries in the matrix.

Find

(*a*) The probability of both A and B being at large 4 minutes later, and

(*b*) The probability that both A and B will be caught eventually.

(*c*) The probability that after 5 minutes A escapes.

We first partition the matrix as follows,

$$A = \left(\begin{array}{ccc|c} \frac{3}{8} & \frac{1}{4} & \frac{1}{4} & \frac{1}{8} \\ \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{2} \\ \frac{1}{4} & 0 & \frac{1}{2} & \frac{1}{4} \\ \hline 0 & 0 & 0 & 1 \end{array}\right) = \begin{pmatrix} P & Q \\ O & I \end{pmatrix} \text{ say}$$

We now find P^n.

The eigenvalues of P are found to be 0, $\frac{3}{4}$, $\frac{3}{8}$, with corresponding eigenvectors

$$\begin{pmatrix} -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \end{pmatrix} = U$$

Since

$$P^n = UD^nU^{-1}$$

$$= \begin{pmatrix} -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & (\frac{3}{4})^n & 0 \\ 0 & 0 & (\frac{3}{8})^n \end{pmatrix} \begin{pmatrix} -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \end{pmatrix}$$

$$= \begin{pmatrix} 0 & \frac{2}{3}(\frac{3}{4})^n & \frac{1}{3}(\frac{3}{8})^n \\ 0 & \frac{1}{3}(\frac{3}{4})^n & \frac{2}{3}(\frac{3}{8})^n \\ 0 & \frac{2}{3}(\frac{3}{4})^n & -\frac{2}{3}(\frac{3}{8})^n \end{pmatrix} \begin{pmatrix} -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \end{pmatrix}$$

$$= \begin{pmatrix} (\frac{2}{3})^2(\frac{3}{4})^n + (\frac{1}{3})^2(\frac{3}{8})^n & \frac{1}{3} \cdot \frac{2}{3}(\frac{3}{4})^n + \frac{1}{3}(\frac{2}{3})(\frac{3}{8})^n & (\frac{2}{3})^2(\frac{3}{4})^n - \frac{2}{3} \cdot \frac{1}{3} \cdot (\frac{3}{8})^n \\ \frac{1}{3} \cdot \frac{2}{3}(\frac{3}{4})^n + \frac{2}{3} \cdot \frac{1}{3} \cdot (\frac{3}{8})^n & (\frac{1}{3})^2 \cdot (\frac{3}{4})^n + (\frac{2}{3})^2(\frac{3}{8})^n & \frac{1}{3} \cdot \frac{2}{3} \cdot (\frac{3}{4})^n - (\frac{2}{3})^2 \cdot (\frac{3}{8})^n \\ (\frac{2}{3})^2(\frac{3}{4})^n - \frac{2}{3} \cdot \frac{1}{3}(\frac{3}{8})^n & \frac{1}{3} \cdot \frac{2}{3} \cdot (\frac{3}{4})^n - (\frac{2}{3})^2(\frac{3}{8})^n & (\frac{2}{3})^2(\frac{3}{4})^n + (\frac{2}{3})^2(\frac{3}{8})^n \end{pmatrix}$$

Now

$$A^n = \begin{pmatrix} P^n & (I - P^n)(I - P)^{-1}.Q \\ O & I \end{pmatrix}$$

$$I - P = \begin{pmatrix} \frac{5}{8} & -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{3}{4} & 0 \\ -\frac{1}{4} & 0 & \frac{1}{2} \end{pmatrix}$$

$$(I-P)^{-1} = \begin{pmatrix} \frac{24}{10} & \frac{8}{10} & \frac{12}{10} \\ \frac{8}{10} & \frac{16}{10} & \frac{4}{10} \\ \frac{12}{10} & \frac{4}{10} & \frac{26}{10} \end{pmatrix} \quad \text{and} \quad Q = \begin{pmatrix} \frac{1}{8} \\ \frac{1}{2} \\ \frac{1}{4} \end{pmatrix}$$

$$(I-P)^{-1}Q = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Now as $n \to \infty$,

$$P^n \to \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

so that

$$\lim_{n\to\infty} A^n = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

i.e. the criminals are caught!

The probability of A and B being at large 4 minutes later is given by the top left hand entry in A^4, i.e.

$$(\tfrac{2}{3})^2(\tfrac{3}{4})^4 + (\tfrac{1}{3})^2(\tfrac{3}{8})^4 = \tfrac{585}{4096} \approx 0{\cdot}1428.$$

The probability that A is at large after 5 minutes is given by the middle entry of the top row of P^n in the case when $n = 5$, i.e.

$$\tfrac{1}{3}(\tfrac{2}{3})(\tfrac{3}{4})^5 + \tfrac{1}{3}(\tfrac{2}{3})(\tfrac{3}{8})^5$$

Exercise 5.3

1. The President of the United States informs A of his decision whether or not to run again in the coming election. A then passes the information to B, B to C and so on. At any stage there is a probability p that the announcement will be changed. We thus have a probability matrix

$$\begin{array}{cc} & & \text{B} \\ & & \begin{array}{cc} \text{Yes} & \text{No} \end{array} \\ \text{A} & \begin{array}{c} \text{Yes} \\ \text{No} \end{array} & \begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix} \end{array}$$

for the transmission of the announcement.

Evaluate the list as $n \to \infty$ for (i) $p = 0{\cdot}1$, (ii) a general value p. What do you notice about the resulting matrix in part (ii)?

2. The storage of a digit in a computer consists of passing it through a number of stages inside the machine. At each stage there is a probability p that the machine will accidentally change the digit, and a

probability q of the machine breaking down. Once the machine has broken down it will remain so. Set up a transition matrix (using Right, Wrong and Broken Down as stages) and evaluate

(i) If $p = 0{\cdot}1$, $q = 0{\cdot}1$ the probability of correct storage in 3 stages.
(ii) What happens as $n \to \infty$ if $p = 0$, $q \neq 0$.
(iii) What happens as $n \to \infty$ if $p \neq 0$, $q = 0$.

For $p = 0{\cdot}1$, $q = 0{\cdot}1$, find the probability that the digit is stored (a) at all, (b) correctly in five stages.

3. If a man earns a high income, there is an 80% chance that his son will also earn a high income; there are 10% chances that he will earn a middle or low income respectively. If he earns a middle income the high:middle:low ratio is 20:60:20%, and for a low income the figures are 25:25:50%. Set up the transition matrix and find the probability that the grandson of a low-income man is a high-income man, and the probability that the great grandson of a high-income man is a low-income man. (N.B. Do not attempt to evaluate eigenvalues: square and cube the transition matrix as it stands.)

4. Assume that in Question 3 the probability of a man having a son is 0·8: construct the new transition matrix and answer the same questions as in Question 3.

5. The theory of genetics as developed by Gregor Mendel postulates that, for any given characteristic, an individual possesses two factors known as genes which determine that characteristic. Each parent passes on one gene to an offspring: the genes of the offspring are independent, and if a parent has two different genes, the one to be transmitted is chosen at random. Each gene may be of two kinds, dominant (denoted by a capital letter) and recessive (denoted by a small one): a person with either two or one dominant genes will show the dominant characteristic, a person with two recessives will show the recessive characteristic. An example is that of height of pea plants: we consider there are two types: tall, dominant (T) and dwarf, recessive (t).

The following matrix shows the probability that a pair of offspring selected at random from those of any one of the pairs of parents on the left exhibit the characteristics shown at the top.

$$\begin{array}{ccc}
 & & \begin{array}{cccccc} \multicolumn{6}{c}{\text{OFFSPRING}} \\ \text{TT Tt} & \text{Tt tt} & \text{TT tt} & \text{Tt Tt} & \text{TT TT} & \text{tt tt} \end{array} \\
\text{PARENTS} & \begin{array}{cc} \text{TT} & \text{Tt} \\ \text{Tt} & \text{tt} \\ \text{TT} & \text{tt} \\ \text{Tt} & \text{Tt} \\ \text{TT} & \text{TT} \\ \text{tt} & \text{tt} \end{array} & \begin{pmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{4} & \frac{1}{4} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{4} & 0 & \frac{1}{4} \\ 0 & 0 & 0 & 1 & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{8} & \frac{1}{4} & \frac{1}{16} & \frac{1}{16} \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}
\end{array}$$

By taking the nth power of this matrix we can thus find the probable characteristics of the nth generation.

Find the probability:

(i) That the initial mating of two hybrids (Tt, Tt) will eventually produce a predominant strain (TT, TT).

(ii) Of ending with a pure recessive strain (tt, tt) if we do not know the original parental types.

Explain why the probability of eventually reaching a pure dominant strain from (TT, tt) is the same as that from (Tt, Tt).

6. In Utopia there are three types of weather and on any particular day the weather belongs to just one of these: 1, sunny; 2, rainy; 3, cloudy but dry. It has been observed that if the weather on a certain day is of type i then that on the following day is of type j with probability P_{ij}, where P_{ij} is the entry in the ith row and jth column of the array below:

$$\begin{array}{ccc} 0{\cdot}3 & 0{\cdot}3 & 0{\cdot}4 \\ 0{\cdot}2 & 0{\cdot}4 & 0{\cdot}4 \\ 0{\cdot}1 & 0{\cdot}3 & 0{\cdot}6 \end{array}$$

If $\mathbf{a} = (a_1, a_2, a_3)$ describes the probabilities of the 3 types of weather on a given day, and $\mathbf{b} = (b_1, b_2, b_3)$ the consequent corresponding probabilities for the next day,

(i) prove that $b_j = \sum_{i=1}^{3} a_i P_{ij}$;

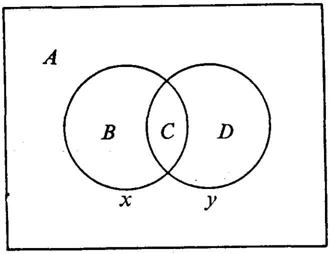
(ii) show that there is an $\mathbf{a}$ such that $\mathbf{b} = \mathbf{a}$ (i.e. there is a *steady state* in which the statistical properties of the weather do no change with time);

(iii) given that the steady state obtains, find the probabilities q_{ij} that, if the weather on a particular day is of type i, then that on the *preceding* day is of type j. (*CS*)

6

Karnaugh Maps and Associated Matrices

6.1 Consider a Venn diagram illustrating two sets, x and y.



The total area can be divided into four pieces corresponding to the four regions A, B, C and D. These regions are given by*

A is $x'y'$	B is $y'x$
C is xy	D is $x'y$

Clearly the whole region is given by

$$x'y' + xy' + x'y + xy$$

These four functions

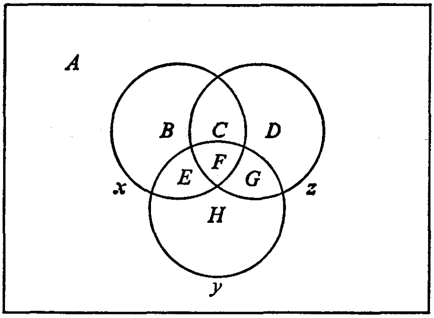
$$xy, \quad xy', \quad x'y \quad \text{and} \quad x'y'$$

are known as **minimal polynomials** (in the case where we are considering two sets).

6.2 Secondly, we can consider a Venn diagram illustrating three sets.

---

* We use xy for $x \cap y$, $x + y$ for $x \cup y$.



This time the area can be divided into eight regions A, B, C, D, E, F, G and H where A is $x'y'z'$, B is $xy'z'$, C is $xy'z$, D is $x'y'z$, E is xyz', F is xyz, G is $x'yz$ and H is $x'yz'$.

This time the whole region is made up of

$$x'y'z' + xy'z' + x'yz' + x'y'z + xy'z + x'yz + xyz$$

Each of these eight functions is a minimum polynomial (in the case where we are considering three sets).

ORDER

The order of the elements in the minimal polynomial is important and found as follows:

For two sets x and y the corresponding J_2* matrix is

$$\begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

This can be thought of as a truth table (or table of closure) for x and y if x is the first row and y the second, i.e.

$$\begin{array}{ccccc} x & 0 & 1 & 0 & 1 \\ y & 0 & 0 & 1 & 1 \end{array}$$

and then considering the columns as the combinations of x and y. For each column we write down the minimal polynomial combination which is true, i.e.

$$\begin{array}{ccccc} x & 0 & 1 & 0 & 1 \\ y & 0 & 0 & 1 & 1 \\ & \uparrow & \uparrow & \uparrow & \uparrow \\ & x'y' & xy' & x'y & xy \end{array}$$

and the order of the functions is shown.

---

* i.e. the matrix with columns in the order of binary numbers reading upwards (see p. 161).

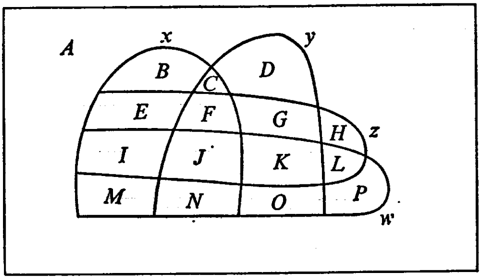
Similarly for the three sets x, y and z, the J_3 matrix is

$$\begin{array}{cc}
\begin{array}{c} x \\ y \\ z \end{array} & \begin{pmatrix} 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} \\
 & \begin{array}{cccccccc} \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow \\ x'y'z' & xy'z' & x'yz' & xyz' & x'y'z & xy'z & x'yz & xyz \end{array}
\end{array}$$

and again we have the order of the functions for the minimal polynomial.

This idea naturally extends to 4 or more sets.

6.3 This idea can be extended to the case in which four sets are considered.



Here there are 16 minimal polynomials of the type $xyzw$ and the other possible combinations.

6.4 The importance of these polynomials is the way in which they can be used to simplify switching circuits. We have met them before when we were given the properties of a switching circuit and were required to find the Boolean function representing the circuit.*

The Venn diagrams which we have drawn are not very useful in their unsystematic outline. To help us in simplifying circuits, a modified version of diagram has been developed by Karnaugh.

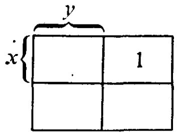
KARNAUGH MAPS (sometimes called VEITCH DIAGRAMS)

6.5 These are rectangular or square representations, rather than the oval or irregular shapes that we have used before.

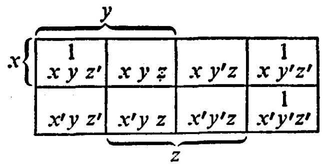
It is not worth using these maps for less than three variables in practice. Here is a map for two variables.

---

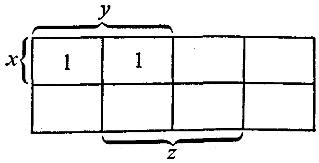
* For example, see *Sets and Logic 2*, Chapter 2.



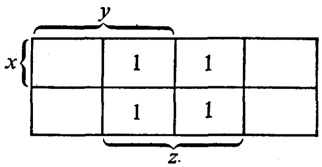
An entry 1 is inserted in the diagram to represent the function to be simplified. The 1 in the top right represents the function xy' and the other squares represent the three other minimal polynomials of two variables. For three variables we have



Here the entries represent $xyz' + x'y'z' + xy'z'$.

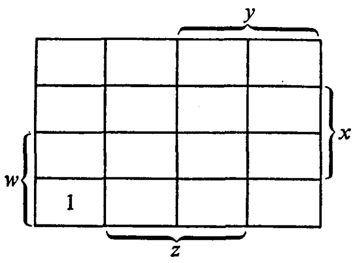


This represents xy (or $xyz' + xyz$).

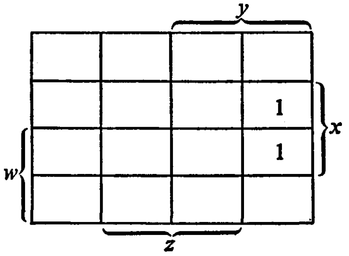


This represents z.

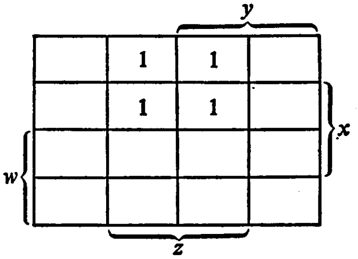
6.6 These maps for three variables are used later on in the matrices of the Karnaugh map (see **6.8**). We now return to the case of four variables to illustrate how the maps may be used to simplify functions.



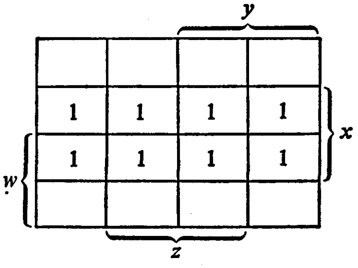
The function illustrated above is $wx'y'z'$. When the function contains all four of its constituents then there is only one entry in the table. When only 3 or 2 or 1 of the constituents is used we have maps as below.



The function xyz'.



The function zw'.



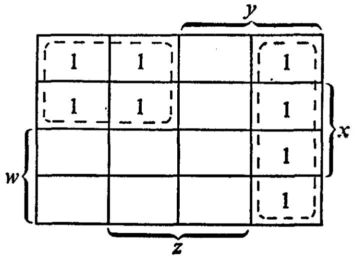
The function x.

Simplification

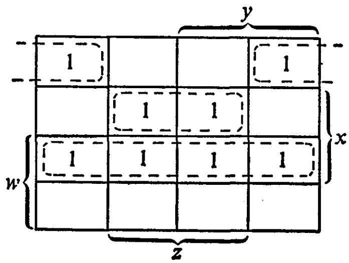
6.7 The digit 1 is placed in each cell required to represent the given function. These digits are then grouped as follows. Marked cells next to each other are grouped into rectangles, squares, rows, columns and pairs allowing 'end-over' adjacencies in both directions until each marked cell is grouped if possible. [Note that you can treat the four corner cells as forming a 'square' because of 'end-over' adjacencies.]

Examples of grouping

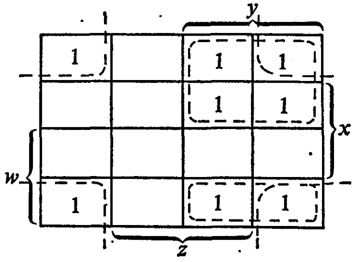
Example 1



Example 2

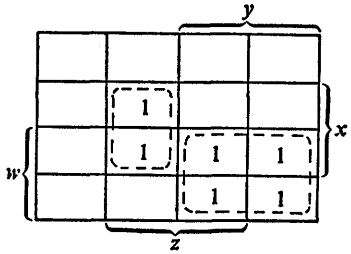


Example 3



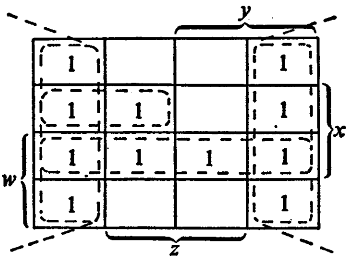
Examples of simplification

Example 1



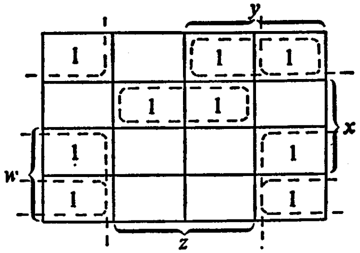
$$f = xyw + x'yz'w + x'yzw + xzy'$$
$$= yw + xzy'$$

Example 2



$$f = xy' + xz' + xyzw + x'y'z' + y'z'w' + yz' + z'w'$$
$$= z' + wx + xy'w' \text{ (or } z' + xy' + wx \text{ by different grouping)}$$

Example 3

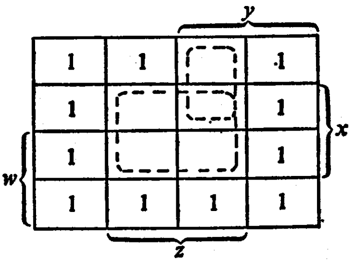


$$f = x'yw' + ywz' + xzw' + x'y'z'w' + xy'zw' + x'yzw'$$
$$= z'w + xzw' + x'yw' + x'z'$$

These are not the only possible groupings in each case. The 'better' the grouping, the simpler will be the final function.

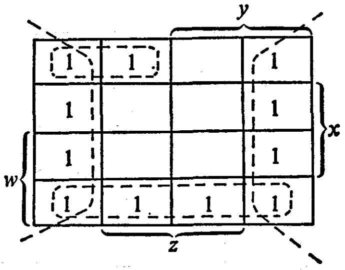
6.8 In some cases it may be much easier to group the empty cells together and hence find the complement of the function. The function can then be found using De Morgan's laws.

Example 4



Here $$f' = xz + yzw'$$

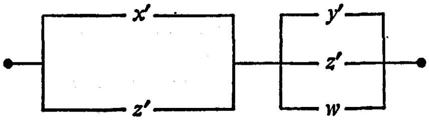
so that $$f = (x' + z') \cdot (y' + z' + w)$$



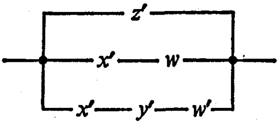
The alternative grouping is given above.

$$f = z' + x'w + x'y'w'$$

The two switching circuits involved are



and



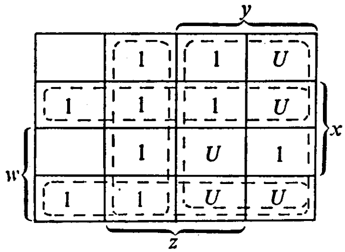
Naturally the first, with the least switches, is the 'better.' There is a better grouping still $(z' + x'(w + y'))$

Exercise 6.1

Use Karnaugh maps to simplify the following functions:

1. $f = wxy'z' + yzw' + wy' + wxy$.
2. $f = wz' + xz' + x'yz' + xyzw + wy' + w'x'y' + w'x'z'$.
3. $f = x'y'w + xyz'w' + xyz' + x'ywz + xywz' + xyz$.
4. $f = x'yzw + xyz'w' + y'zw' + wxz' + w'x'y' + w'y'z$.
5. $f = x'y'w' + x'y'z + x'y'w + zy'w' + xyz'w' + xy'z'w'$.
6. $f = xyzw + xy'z'w + x'y'zw + x'y'z'w' + x'yzw + xy'z'w'$.

It sometimes occurs that we do not mind what the value of one of the four functions may be. This situation can be plotted on the map and we can see the function as best we can in our grouping procedure. Suppose we have for example,

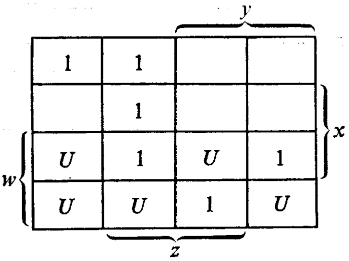


	1	1	U
1	1	1	U
	1	U	1
1	1	U	U

with y unrestricted.

We have entered a U in the map to show that the entry can be either 1 or 0,* as is most convenient for us in the simplification process.

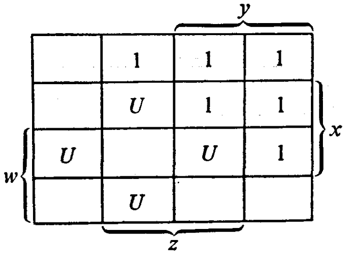
7. Simplify the following and give the function,



1	1		
	1		
U	1	U	1
U	U	1	U

Here w is unrestricted.

8.



	1	1	1
	U	1	1
U		U	1
	U		

Here the unrestricted entries occur due to certain input conditions on the circuit.

Give the simplified function for the above map,

(*a*) by grouping to find the function, and
(*b*) by grouping to find the complementary function, and hence the function using De Morgan's laws.

---

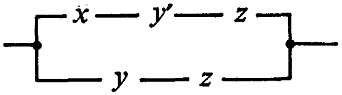
* All the U's must be the same, either all 1 or all 0.

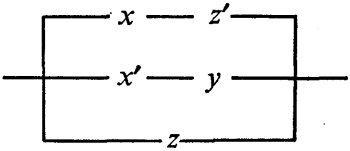
These maps are of little use except in the simplification of functions of *four* variables.

MATRICES ASSOCIATED WITH KARNAUGH MAPS

6.8 A Boolean matrix may be defined as a matrix which contains Boolean entries, i.e. 1 and 0. We shall see that such a matrix can be arrived at in a variety of ways. Combinations of these matrices, too, can be achieved in different ways. Here we consider the matrices associated with the Karnaugh maps.

6.9 Consider two switching circuits of three variables.





The Boolean functions which represent these two circuits are

$$f_1 = xy'z + yz$$

and

$$f_2 = xz' + x'y + z.$$

These two can be written in the form of minimal polynomials by the use of Karnaugh maps as shown below.

$$\begin{array}{c|c|c|c|c|}
 & \multicolumn{2}{c}{\overbrace{\qquad\qquad}^{y}} & \multicolumn{2}{c}{} \\ \cline{2-5}
x\{ & 0 & 1 & 1 & 0 \\ \cline{2-5}
 & 0 & 1 & 0 & 0 \\ \cline{2-5}
 & \multicolumn{1}{c}{} & \multicolumn{2}{c}{\underbrace{\qquad\qquad}_{z}} & \multicolumn{1}{c}{}
\end{array}$$

The function f_1.

$$\begin{array}{c|c|c|c|c|}
 & \multicolumn{2}{c}{\overbrace{\qquad\qquad}^{y}} & \multicolumn{2}{c}{} \\ \cline{2-5}
x\{ & 1 & 1 & 1 & 1 \\ \cline{2-5}
 & 1 & 1 & 1 & 0 \\ \cline{2-5}
 & \multicolumn{1}{c}{} & \multicolumn{2}{c}{\underbrace{\qquad\qquad}_{z}} & \multicolumn{1}{c}{}
\end{array}$$

The function f_2.

f_1 can then be expressed as the matrix

$$\begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

and f_2 as

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

6.10 When the two circuits are placed in *series* the overall transmission function is given by

$$f = f_1 f_2$$

where $f_1 f_2$ is found by multiplying together the corresponding elements of f_1 and f_2. Here

$$f_1 . f_2 = \begin{pmatrix} 0.1 & 1.1 & 1.1 & 0.1 \\ 0.1 & 1.1 & 0.1 & 0.0 \end{pmatrix}$$

i.e.
$$f = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

where we are using the laws of Boolean algebra, where $0.0 = 0$; $0.1 = 1.0 = 0$; $1.1 = 1$. On the other hand, if the two circuits were placed in *parallel*, the overall transmission function would be given by

$$g = f_1 + f_2$$

where the elements of g are found by adding together the corresponding elements of f_1 and f_2, i.e.

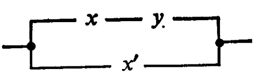
$$f_1 + f_2 = \begin{pmatrix} 0+1 & 1+1 & 1+1 & 0+1 \\ 0+1 & 1+1 & 0+1 & 0+0 \end{pmatrix}$$

i.e.
$$g = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}.$$

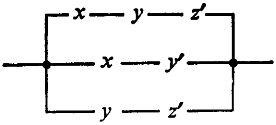
The relevant laws of Boolean algebra here are

$$0 + 0 = 0; \quad 0 + 1 = 1 + 0 = 1; \quad 1 + 1 = 1.$$

6.11 It is possible to combine circuits containing functions of a different number of variables by expanding the Karnaugh map to the size required for the circuit containing the most variables. For example, consider the two circuits given below:

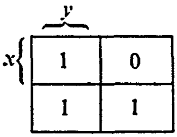


$f_1 = xy + x'$



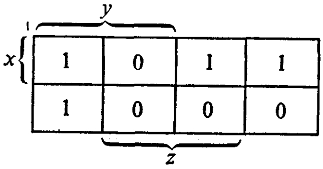
$f_2 = xyz' + xy' + yz'$

Here the first circuit is a function of two variables (x and y) and the second is a function of three variables. The two associated maps are



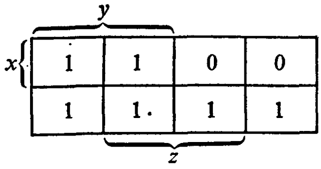
1	0
1	1

and



1	0	1	1
1	0	0	0

In order to combine these two, the first map is expanded to three variables to give:



1	1	0	0
1	1	1	1

and the two transmission matrices are now

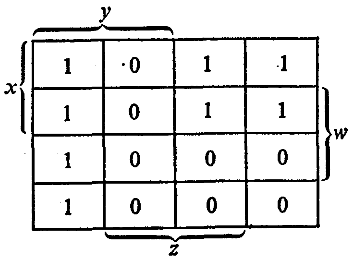
$$\begin{pmatrix}1 & 1 & 0 & 0\\ 1 & 1 & 1 & 1\end{pmatrix} \text{ and } \begin{pmatrix}1 & 0 & 1 & 1\\ 1 & 0 & 0 & 0\end{pmatrix}$$

which can now be dealt with as required.

6.12 The same procedure can be carried out when the circuits contain, for instance, functions of three variables and functions of four variables.

Example

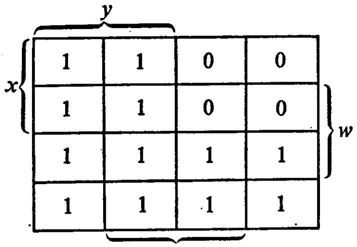
Suppose we wished to combine our second switching circuit with another containing four variables. The Karnaugh map expanded would be



1	0	1	1
1	0	1	1
1	0	0	0
1	0	0	0

$$f_2 = xyz' + xy' + yz'$$

Similarly if our first circuit were to be combined with another of four variables we should have:



1	1	0	0
1	1	0	0
1	1	1	1
1	1	1	1

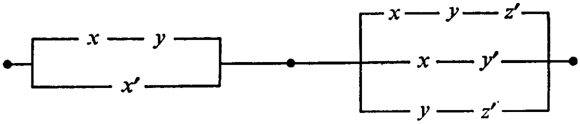
$f_1 = xy + x'.$

The matrix representation here would be

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

6.13 We now attempt to demonstrate that multiplication and addition of these matrices as defined does give the required transmission matrix.

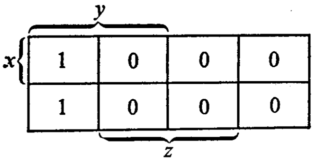
Consider the two switching circuits in **6.11**, put in series, i.e.



$f_1 = xy + x'$ $\qquad$ $f_2 = xyz' + xy' + yz'.$

We already know* that when circuits are taken in series the resulting function is f_1f_2. We can then consider f_1f_2 and see if this gives the required matrix of minimal polynomials.

$$\begin{aligned} f_1f_2 &= (xy + x')(xyz' + xy' + yz') \\ &= (xy + x')(yz' + xy') \\ &= xyz' + x'yz'. \end{aligned}$$



1	0	0	0
1	0	0	0

The map of f_1f_2.

---

* cf. *Sets and Logic 2.*

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

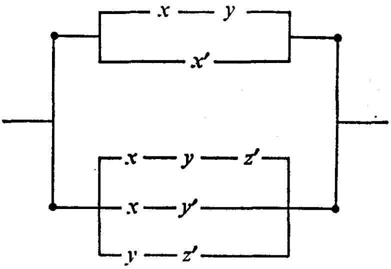
The transmission matrix $f_1 f_2$.

However, using the defined law of multiplication we have

$$f_1 f_2 = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$
$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \text{ as before.}$$

This is not a proof of the method, but an illustration that this method does give us the desired result.

6.14 As a second illustration consider these two switching circuits now in parallel, i.e.

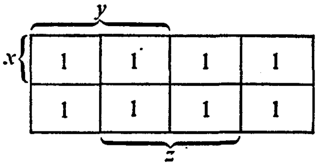


$$f_1 = xy + x'$$
$$f_2 = xyz' + xy' + yz'.$$

We have seen before* that circuits in parallel are combined by addition to give the function representing the circuit. Here

$$\begin{aligned} f_1 + f_2 &= xy + x' + xyz' + xy' + yz' \\ &= xy + x' + yz' + xy' \\ &= x + x' + yz' \\ &= 1. \end{aligned}$$

The map of $f_1 + f_2$ is thus



1	1	1	1
1	1	1	1

---

* *Sets and Logic 2.*

with matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

On the other hand, combining the two matrices for f_1 and f_2 gives:

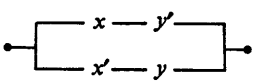
$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

as before. Note that this circuit is always ‘on’.

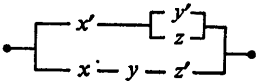
Notice that matrices of this type are commutative since Boolean Algebra elements are commutative under both addition and multiplication.

Exercise 6.2

1.



s_1



s_2

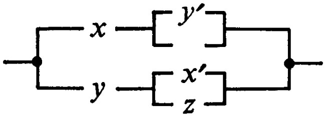
Find the transmission matrices of

(*a*) S_1, (*b*) S_2, (*c*) S_1S_2, (*d*) $S_1 + S_2$.

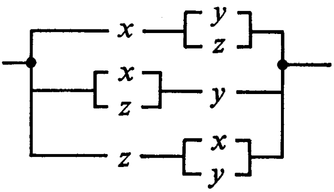
Draw simplified circuits for (*c*) and (*d*).

2. Find the transmission matrices of

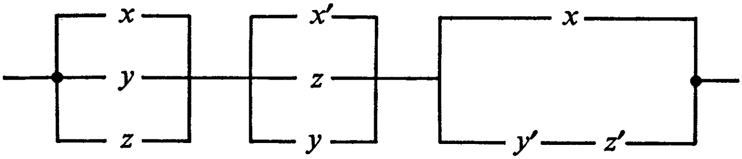
(*a*) p, (*b*) q, (*c*) r, (*d*) pq, (*e*) $p + r$, (*f*) $p + q + r$, (*g*) pqr.



p

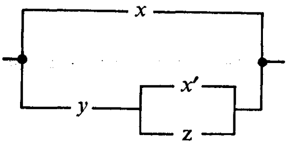


q

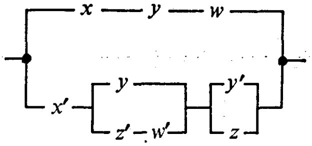


r

3. Find the transmission matrices of



t_1



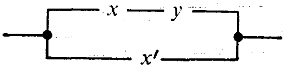
t_2

(*a*) t_1, (*b*) t_2, (*c*) t_1t_2, (*d*) $t_1 + t_2$.

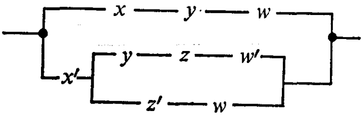
Draw a simplified circuit for (*c*).

4. Find the transmission matrices of

(*a*) u_1, (*b*) u_2, (*c*) $u_1 + u_2$, (*d*) u_1u_2.



u_1



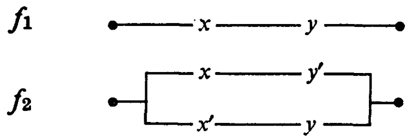
u_2

7

Boolean Matrices

7.1 As we have already mentioned, matrices with Boolean entries can be formed in a number of ways. Consider the two functions

$$f_1 = xy$$
$$f_2 = xy' + yx'$$



These can be expressed in terms of their minimal polynomials as

$$f_1 = 0\,.\,x'y' + 0\,.\,x'y + 0\,.\,xy' + 1\,.\,xy$$
$$f_2 = 0\,.\,x'y' + 1\,.\,x'y + 1\,.\,xy' + 0\,.\,xy$$

and these equations themselves may be represented in matrix form as

$$\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix} . \begin{pmatrix} x \\ y \end{pmatrix}$$

This can be written

$$\mathbf{f} = A\mathbf{z}$$

where A is the matrix of the coefficients and **f** and **z** are both vectors.

J matrices

7.2 These matrices are the Boolean matrices having their entries in the order of binary numbers (reading from the bottom upwards). For instance, the first three J matrices are

$$J_1 = (0 \quad 1) \qquad J_2 = \begin{array}{c} \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \\ \begin{array}{cccc} \uparrow & \uparrow & \uparrow & \uparrow \\ 00 & 01 & 10 & 11 \end{array} \end{array}$$

$$\begin{array}{cc} J_3 = & \left(\begin{array}{cccccccc} 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{array}\right) \\ & \begin{array}{cccccccc} \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow \\ 000 & 001 & 010 & 011 & 100 & 101 & 110 & 111 \end{array} \end{array}$$

Consider now any column matrix or vector $\mathbf{z}$. Each vector $\mathbf{z}$ will be equal to one and only one of the columns in the corresponding J matrix.

For example, the vector

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

is equal to the *third* column of J_2; whilst the vector

$$\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

is equal to the *seventh* column of J_3.

7.3 We now define a special law for Boolean matrix multiplication. Suppose we have

$$A \,.\, B = C$$

then the columns of C are obtained by comparing the columns of B with the corresponding J matrix—suppose the column of B is the second in the J matrix then the column of C chosen is the second in A.

For example to find

$$\underset{A}{\begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix}} \cdot \underset{B}{\begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 \end{pmatrix}}$$

The corresponding J matrix is

$$\begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

i.e. J_2.

The first column of B $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ is the third column in J_2, so the first column of C is the *third* column of A, i.e. $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Similarly the second column of B is $\begin{pmatrix}0\\0\end{pmatrix}$ which is the first column of J_2; hence the second column of C is the *first* column of A, i.e. $\begin{pmatrix}1\\1\end{pmatrix}$, and so on. Here

$$\begin{pmatrix}1 & 0 & 1 & 0\\1 & 1 & 0 & 1\end{pmatrix}.\begin{pmatrix}0 & 0 & 1 & 0\\1 & 0 & 1 & 1\end{pmatrix}=\begin{pmatrix}1 & 1 & 0 & 1\\0 & 1 & 1 & 0\end{pmatrix}$$

7.4 It is interesting to see whether matrices of this type commute under multiplication.

$$\begin{pmatrix}0 & 0 & 1 & 0\\1 & 0 & 1 & 1\end{pmatrix}.\begin{pmatrix}1 & 0 & 1 & 0\\1 & 1 & 0 & 1\end{pmatrix}=\begin{pmatrix}0 & 1 & 0 & 1\\1 & 1 & 0 & 1\end{pmatrix}$$

So that

$$A.B \neq B.A$$

(Note that $AJ = A = JA$.)

Boolean matrices which have two or more of their columns equal are said to be *singular*. The inverse of a Boolean matrix can only be found (in fact only exists) when the matrix is non-singular.

7.5 Since the J matrices are the identity matrices we can find the inverse of a matrix by reversing the process of multiplication

$$A.A^{-1} = J$$

To find the inverse of

$$A = \begin{pmatrix}1 & 1 & 0 & 0\\0 & 1 & 0 & 1\end{pmatrix}$$

i.e. so that

$$\begin{pmatrix}1 & 1 & 0 & 0\\0 & 1 & 0 & 1\end{pmatrix}.\begin{pmatrix} & & & \\ & & & \end{pmatrix}=\begin{pmatrix}0 & 1 & 0 & 1\\0 & 0 & 1 & 1\end{pmatrix}$$

the first column of A^{-1} is found as follows. The first column of J is the *third* column of A—so the first column of A^{-1} is $\begin{pmatrix}0\\1\end{pmatrix}$, the 3rd column of J.

The second column of A^{-1} is found in the same way. The second column of J is *first* column of A so the second column of A^{-1} is $\begin{pmatrix}0\\0\end{pmatrix}$, the first column of J, and so on. So

$$A^{-1} = \begin{pmatrix}0 & 0 & 1 & 1\\1 & 0 & 1 & 0\end{pmatrix}$$

Also

$$\begin{pmatrix}0&0&1&1\\1&0&1&0\end{pmatrix}.\begin{pmatrix}1&1&0&0\\0&1&0&1\end{pmatrix}=\begin{pmatrix}0&1&0&1\\0&0&1&1\end{pmatrix}$$

so that

$$A^{-1}.A=A.A^{-1}=J$$

Solution of Boolean matrix equations

7.6 Consider the equation

$$\mathbf{f}=A\mathbf{z}$$

where A and $\mathbf{z}$ are known and we wish to find $\mathbf{f}$.

For example, suppose

$$\begin{pmatrix}f_1\\f_2\end{pmatrix}=\begin{pmatrix}1&0&1&0\\1&1&0&1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}$$

and we wish to find $\begin{pmatrix}f_1\\f_2\end{pmatrix}$. The process is carried out as follows: $\mathbf{z}$ is the *second* column of J_2 so that $\mathbf{f}$ is the *second* column of A, i.e. $\begin{pmatrix}0\\1\end{pmatrix}$. As a second example consider

$$\begin{pmatrix}f_1\\f_2\\f_3\end{pmatrix}=\begin{pmatrix}0&1&1&0&0&1&0&1\\1&0&1&0&1&1&0&1\\0&0&0&0&1&1&1&0\end{pmatrix}.\begin{pmatrix}1\\1\\0\end{pmatrix}$$

The vector

$$\begin{pmatrix}1\\1\\0\end{pmatrix}$$

is the *fourth* column of J_3, so

$$\begin{pmatrix}f_1\\f_2\\f_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}$$

the *fourth* column of the matrix above.

7.7 Eigenvectors occur sometimes, i.e. when

$$A\mathbf{z}=\mathbf{z}$$

For example $\begin{pmatrix}1\\1\end{pmatrix}$ is an eigenvector of

$$\begin{pmatrix}1&0&0&1\\0&1&0&1\end{pmatrix}$$

since

$$\begin{pmatrix}1 & 0 & 0 & 1\\0 & 1 & 0 & 1\end{pmatrix}\cdot\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}1\\1\end{pmatrix}$$

It is possible to have more than one eigenvector of a matrix. For example, consider

$$\begin{pmatrix}0 & 1 & 0 & 1\\1 & 0 & 0 & 1\end{pmatrix}$$

This has two eigenvectors

$$\begin{pmatrix}1\\0\end{pmatrix} \text{ and } \begin{pmatrix}1\\1\end{pmatrix}$$

since

$$\begin{pmatrix}0 & 1 & 0 & 1\\1 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}1\\0\end{pmatrix} \text{ and } \begin{pmatrix}0 & 1 & 0 & 1\\1 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}1\\1\end{pmatrix}$$

How many eigenvectors has J_n?

Exercise 7.1

1. Multiply the following Boolean matrices:

(*a*) $\begin{pmatrix}1 & 1 & 0 & 1\\0 & 1 & 0 & 0\end{pmatrix}\cdot\begin{pmatrix}0 & 1 & 0 & 1\\1 & 1 & 0 & 1\end{pmatrix}$

(*b*) $\begin{pmatrix}0 & 1 & 0 & 1\\1 & 1 & 0 & 1\end{pmatrix}\cdot\begin{pmatrix}1 & 1 & 0 & 1\\0 & 1 & 0 & 0\end{pmatrix}$

2. Find the product

$$\begin{pmatrix}0 & 1 & 1 & 0 & 0 & 1 & 1 & 0\\1 & 1 & 1 & 1 & 0 & 0 & 1 & 0\\0 & 0 & 1 & 0 & 1 & 0 & 0 & 0\end{pmatrix}\cdot\begin{pmatrix}0 & 1 & 0 & 1 & 0 & 1 & 0 & 0\\0 & 0 & 1 & 0 & 1 & 1 & 0 & 1\\1 & 0 & 1 & 1 & 0 & 1 & 0 & 0\end{pmatrix}$$

3. Are Boolean matrices associative under multiplication of its type? Check using the three matrices

$$\begin{pmatrix}1 & 0 & 0 & 1\\1 & 1 & 0 & 0\end{pmatrix}, \begin{pmatrix}0 & 0 & 0 & 1\\1 & 1 & 1 & 1\end{pmatrix}, \begin{pmatrix}1 & 0 & 1 & 1\\0 & 1 & 1 & 0\end{pmatrix}$$

4. Find the inverse matrix of

$$(a)\ \begin{pmatrix}1 & 0 & 1 & 0\\1 & 1 & 0 & 0\end{pmatrix} \text{ and } (b)\ \begin{pmatrix}0 & 1 & 1 & 0\\0 & 1 & 0 & 1\end{pmatrix}$$

5. Find the inverse matrix of

$$\begin{pmatrix}0 & 1 & 0 & 1 & 1 & 0 & 1 & 0\\1 & 1 & 0 & 1 & 0 & 0 & 0 & 1\\0 & 0 & 1 & 1 & 1 & 0 & 0 & 1\end{pmatrix}$$

6. Find the eigenvectors of

(*a*) $\begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}$ (*b*) $\begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix}$

(*c*) $\begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \end{pmatrix}$ (*d*) $\begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$

(*e*) $\begin{pmatrix} 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$ (*f*) $\begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \end{pmatrix}$

7. Solve the following equations

(*a*) $\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

(*b*) $\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

8. Solve the equations

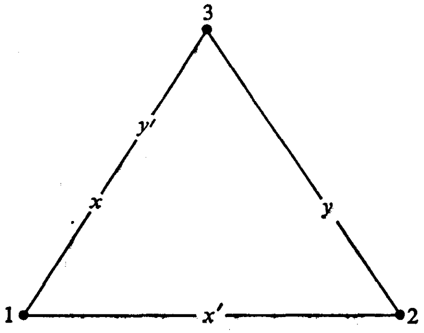
(*a*) $\begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$

(*b*) $\begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$

Output vectors from switching circuits

7.8 The output from a circuit across a pair of terminals, for given input conditions can be found as a vector.

Consider the circuit



This gives

$$\begin{aligned} f_{12} &= xy' + x'y \\ f_{13} &= x' + xy'y = x' \\ f_{23} &= y + x'xy' = y \end{aligned}$$

We can now find each of these outputs in terms of the minimal polynomials, by the use of Karnaugh maps.

$$f_{12}\ \begin{array}{c|c|c|} \multicolumn{1}{c}{} & \multicolumn{1}{c}{\overbrace{\hphantom{0}}^{y}} & \multicolumn{1}{c}{} \\ \cline{2-3} x\{ & 0 & 1 \\ \cline{2-3} & 1 & 0 \\ \cline{2-3} \end{array} \qquad f_{13}\ \begin{array}{c|c|c|} \multicolumn{1}{c}{} & \multicolumn{1}{c}{\overbrace{\hphantom{0}}^{y}} & \multicolumn{1}{c}{} \\ \cline{2-3} x\{ & 0 & 0 \\ \cline{2-3} & 1 & 1 \\ \cline{2-3} \end{array} \qquad f_{23}\ \begin{array}{c|c|c|} \multicolumn{1}{c}{} & \multicolumn{1}{c}{\overbrace{\hphantom{0}}^{y}} & \multicolumn{1}{c}{} \\ \cline{2-3} x\{ & 1 & 0 \\ \cline{2-3} & 1 & 0 \\ \cline{2-3} \end{array}$$

These can be expressed as:

$$\begin{aligned} f_{12} &= 0x'y' + 1xy' + 1x'y + 0xy \\ f_{13} &= 1x'y' + 0xy' + 1x'y + 0xy \\ f_{23} &= 0x'y' + 0xy' + 1x'y + 1xy \end{aligned}$$

7.9 This enables us to find the output vector from a circuit if we are given the input vector.

In Boolean matrix form we have

$$\begin{pmatrix} f_{12} \\ f_{13} \\ f_{23} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix}$$

As the matrix $\begin{pmatrix} x \\ y \end{pmatrix}$ consists of two variables the relevant J matrix is J_2, i.e.

$$\begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

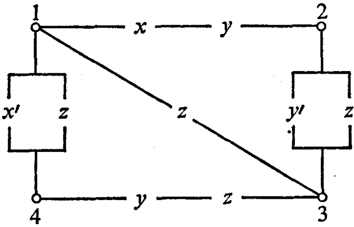
To solve: Suppose the input vector is $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$; this is equal to the first column in our J matrix, so that the output will be the first column of our matrix, i.e.

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

Similarly, if we consider an input vector of $\begin{pmatrix}0\\1\end{pmatrix}$ this is equal to the third column of our J matrix and hence the output vector is

$$\begin{pmatrix}1\\1\\1\end{pmatrix}.$$

7.10 Consider as a second example the circuit below:

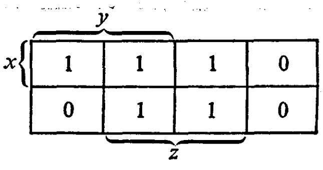


$$\begin{aligned}
f_{12} &= xy + z(y' + z') + (x' + z)(yz)(y' + z')\\
f_{13} &= z + xy(y' + z') + (x' + z)yz\\
f_{14} &= x' + z + xy(y' + z')(yz) + z(yz)\\
f_{23} &= y' + z' + xyz + xy(x' + z)(yz)\\
f_{24} &= xy(x' + z) + (y' + z')yz + (y' + z')z(x' + z) + xyzyz\\
f_{34} &= yz + z(x' + z) + (y' + z')xy(x' + z)
\end{aligned}$$

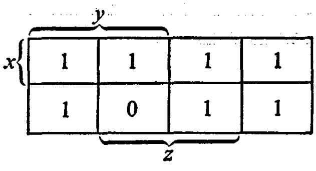
These output functions can, first of all, be simplified using the laws of Boolean algebra, to give:

$$\begin{aligned}
f_{12} &= xy + zy'\\
f_{13} &= z + xyz'\\
f_{14} &= x' + z\\
f_{23} &= y' + z' + xyz\\
f_{24} &= xyz + y'z\\
f_{34} &= z
\end{aligned}$$

7.11 The next step is to find the Karnaugh map of each function and eventually to write the functions in terms of the minimal polynomials. The Karnaugh maps of f_{13} and f_{23} are:



x	1	1	1	0
	0	1	1	0



x	1	1	1	1
	1	0	1	1

These give us, in terms of the minimal polynomial

$$f_{13} = 0x'y'z' + 0xy'z' + 0x'yz' + 1xyz' + 1x'y'z + 1xy'z + 1x'yz + 1xyz$$

$$f_{23} = 1x'y'z' + 1xy'z' + 1x'yz' + 1xyz' + 1x'y'z + 1xy'z + 0x'yz + 1xyz$$

The whole set of six equations can thus be written in matrix form as:

$$\begin{pmatrix} f_{12} \\ f_{13} \\ f_{14} \\ f_{23} \\ f_{24} \\ f_{34} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

As previously, we can solve these equations for any given input vector

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

The appropriate J matrix is J_3, i.e.

$$\begin{pmatrix} 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Suppose the input is

$$\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix};$$

then the output will be the fourth column of our matrix, i.e.

$$\begin{pmatrix} f_{12} \\ f_{13} \\ f_{14} \\ f_{23} \\ f_{24} \\ f_{34} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

7.12 Our next problem is to investigate the processes described, i.e. the multiplication of these matrices and the problem of finding the output from a circuit with given input. Consider again, the example on p. 164,

$$\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

and return to the original equations, i.e.

$$f_1 = 1x'y' + 0xy' + 1x'y + 0xy$$
$$f_2 = 1x'y' + 1xy' + 0x'y + 1xy$$

The vector $\begin{pmatrix}1\\0\end{pmatrix}$ tells us that $x = 1$, $y = 0$, i.e. the only non-zero element in the minimal polynomial is the one containing xy'—corresponding to the second column. Alternatively, if

$$f_1 = x'y' + x'y \quad \text{and} \quad x = 1, y = 0$$

we have

$$f_1 = 0$$

similarly when

$$f_2 = x'y' + xy' + xy \quad \text{and} \quad x = 1, y = 0$$

we have

$$f_2 = 1$$

i.e.

$$\begin{pmatrix}f_1\\f_2\end{pmatrix} = \begin{pmatrix}0\\1\end{pmatrix},$$

the second column of the matrix as before.

7.13 Multiplication of these matrices is carried out likewise; by considering, for example

$$\begin{pmatrix}1 & 0 & 1 & 0\\1 & 1 & 0 & 1\end{pmatrix} \cdot \begin{pmatrix}1 & 1 & 0 & 0\\0 & 1 & 0 & 1\end{pmatrix}$$

as being constructed from

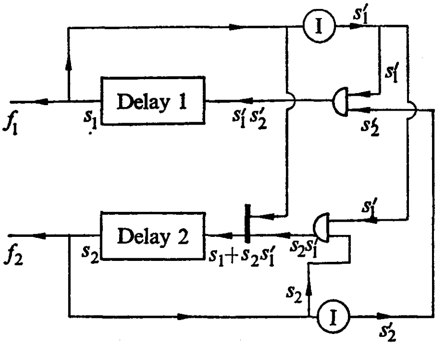
$$\underset{\textit{first column}}{\begin{pmatrix}1 & 0 & 1 & 0\\1 & 1 & 0 & 1\end{pmatrix} \cdot \begin{pmatrix}1\\0\end{pmatrix}}, \qquad \underset{\textit{second column}}{\begin{pmatrix}1 & 0 & 1 & 0\\1 & 1 & 0 & 1\end{pmatrix} \cdot \begin{pmatrix}1\\1\end{pmatrix}}$$

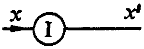
$$\underset{\textit{third column}}{\begin{pmatrix}1 & 0 & 1 & 0\\1 & 1 & 0 & 1\end{pmatrix} \cdot \begin{pmatrix}0\\0\end{pmatrix}} \quad \text{and} \quad \underset{\textit{fourth column}}{\begin{pmatrix}1 & 0 & 1 & 0\\1 & 1 & 0 & 1\end{pmatrix} \cdot \begin{pmatrix}0\\1\end{pmatrix}}$$

Circuits involving delay units

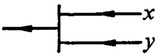
7.14 Suppose that $S_1, S_2, \ldots S_n$ represent the states of n delay units at any specified instant of time. Then the outputs $f_1, f_2, \ldots f_n$ represent the states of these units 1 instant later.

The diagram below shows how a simple delay circuit is built up.*



Where  represents the Inverter

$x.y$ AND Gate

and $x + y$  OR Gate

In this case

$$\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} S_1 \\ S_2 \end{pmatrix}$$

(the actual equations are $f_1 = S_1'S_2'$, $f_2 = S_1 + S_1'S_2$).

7.15 Suppose that at time instant zero, S_1 and S_2 are both zero, i.e.

$$\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

so

$$\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

This output is then fed into the circuit again as an input to give a new output, i.e.

$$\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

---

* This is identical to the usual method of making gated circuits—see *Sets and Logic 2*.

i.e.

$$\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

and this is fed in again to give

$$\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

i.e.

$$\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

and so on.

f_1	f_2	*Time instant*
0	0	0
1	0	1
0	1	2
0	1	3

Here the output clearly stabilizes as $f_1 = 0, f_2 = 1$ after 2 time instants.

If the initial conditions on S_1 and S_2 were different we still achieve stabilization quickly, i.e.

If $$S_1 = 0, \quad S_2 = 1, \quad f_1 = 0, \quad f_2 = 1$$

at once, and similarly when

$$S_1 = 1, \quad S_2 = 0 \quad \text{or} \quad S_1 = 1, \quad S_2 = 1$$

No matter what the initial states of the delay units are, the outputs lock on to $f_1 = 0, f_2 = 1$ after, at most two instants of time.

Exercise 7.2

Draw the delay circuits represented by the following equations. Investigate the output cycles.

1. $f_1 = S_1S_2 + S_1'S_2'$
   $f_2 = S_2'$.
2. $f_1 = S_2 + S_3'$
   $f_2 = S_1(S_2 + S_3')$
   $f_3 = S_1'S_3$.
3. $f_1 = S_1S_2S_3' + S_1S_2'$
   $f_2 = S_2(S_2S_3 + S_1)$
   $f_3 = S_1'$.
4. $f_1 = S_1S_2' + S_1'S_2$
   $f_2 = S_1'$.

Notice that a locking state occurs whenever the vector $\begin{pmatrix} f_1 \\ f_2 \end{pmatrix}$ is an eigenvector of the equation. In the example given

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

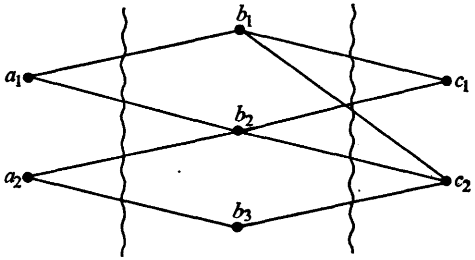
i.e. $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ is an eigenvector of the equation. What happens if we consider every other time instant? Clearly it makes no difference if the system has a locked output.

8

Matrices and Networks

8.1 Matrices provide an easy method of describing networks. The following methods are applicable when a computer can be used. The instructions are simplified to give the ideas.

8.2 Consider the following network of roads linking two towns in county *A*, three towns in county *B* and two towns in county *C*.



We can describe the roads connecting towns in counties *A* and *B* as follows:

	b_1	b_2	b_3
a_1	1	1	0
a_2	0	1	1

In usual matrix notation we may write

$$P = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

The matrix describing the connections between the towns in county *B* and in county *C* is

	c_1	c_2
b_1	1	1
b_2	1	1
b_3	0	1

or in usual notation

$$Q = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 1 \end{pmatrix}$$

Now consider a combination of P and Q (we call it $P\,.\,Q$) and see what logical interpretation we might give to the resulting matrix.

$$P\,.\,Q = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 1 & 2 \end{pmatrix}$$

This final matrix must tell us something about the roads from the towns in county A to the towns in county C. Consider

	c_1	c_2
a_1	2	2
a_2	1	2

This tells us that there are *two* routes from a_1 to c_1, *two* routes from a_1 to c_2 (via b_1 or b_2) and so on. This is the information given by the product matrix.

This method can be extended indefinitely to give us the number of routes from any number of towns to any other numbers of towns.

Some problems

8.3 In this representation, what would be the identity matrix? An inverse matrix? Clearly matrices with negative entries can have no physical meaning in the situation we are describing. The idea of an inverse matrix can be given a meaning. The matrix P represents the routes from towns in county A to towns in county B; the 'inverse' matrix will describe the routes from town in county B to towns in county A, i.e. here

$$P = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

so that the inverse of P

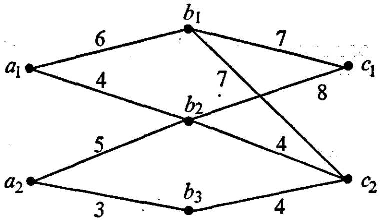
$$P^{-1} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix}$$

or the inverse of P

$$P^{-1} = P' \qquad \text{(the transpose of } P\text{).}$$

8.4 What happens if we multiply P by P'? What information is obtained from the product?

Consider again the original network



The numbers inserted represent the distance in miles between the towns.

We can now make up a distance matrix, where each entry represents the distance between the towns. Here

$$P = \begin{array}{c|ccc} & b_1 & b_2 & b_3 \\ \hline a_1 & 6 & 4 & - \\ a_2 & - & 5 & 3 \end{array} \qquad Q = \begin{array}{c|cc} & c_1 & c_2 \\ \hline b_1 & 7 & 7 \\ b_2 & 8 & 4 \\ b_3 & - & 4 \end{array}$$

The – shows that the towns are not directly linked. Here the – would be replaced by ∞ if a computer were used. In practice a very large number, say 10^6 would probably be used.

8.5 How long are the *shortest* routes between towns in county A and towns in county C? We can work out the entries by trial and error. They are:

$$R = \begin{array}{c|cc} & c_1 & c_2 \\ \hline a_1 & 12 & 8 \\ a_2 & 13 & 7 \end{array}$$

The problem is—how do we combine P and Q to obtain R. The solution is easier to obtain than to explain. Consider matrix multiplication of P and Q. We should consider first

$$(6 \quad 4 \quad -) \begin{pmatrix} 7 \\ 8 \\ - \end{pmatrix}$$

In this case, instead of multiplying corresponding entries and adding, we *add* corresponding entries and write down the *minimum*,

e.g. R is

$$\begin{pmatrix} \textit{min} \text{ of } \{6+7, 4+8, - + -\} & \textit{min} \text{ of } \{6+7, 4+4, - + 4\} \\ \textit{min} \text{ of } \{- + 7, 5+8, 3 + -\} & \textit{min} \text{ of } \{- + 7, 4+5, 3+4\} \end{pmatrix}$$

or

$$\begin{pmatrix} 12 & 8 \\ 13 & 7 \end{pmatrix}$$

We call this operation MIN, hence R = P MIN Q. Note that the dashes rule out the entries where they occur.

8.6 Notice that similar situations arise where the numbers on each road represent, for example:

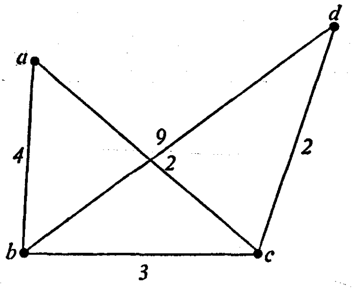
1. The average time taken for each stretch, or
2. The capacity of drains.

In the first example the process to find the shortest time is identical to that shown for finding the shortest distance. In the second example,* the capacity of the drain from a to b is 6 and the capacity for b to c is 7, so that the maximum capacity from a to c is 6. In this case our maximum drain matrix would be:

$$D = \begin{array}{c|cc} & c_1 & c_2 \\ \hline a_1 & 6 & 6 \\ a_2 & 5 & 4 \end{array}$$

Give a rule for finding D from P and Q. Investigate this rule. For example, is it commutative or associative?

8.7 Consider a very simple network with distances inserted



---

* Drain flow in two ways? Investigate.

The connecting matrix with distances inserted is then

$$S_1 = \begin{array}{c|cccc} & a & b & c & d \\ \hline a & 0 & 4 & 2 & - \\ b & 4 & 0 & 3 & 9 \\ c & 2 & 3 & 0 & 2 \\ d & - & 9 & 2 & 0 \end{array}$$

8.8 This contains information about the *one-link* routes. What are the shortest routes between *a*, *b*, *c* and *d* if we allow *one-link* or *two-link* routes?

$$S_2 = \begin{array}{c|cccc} & a & b & c & d \\ \hline a & 0 & 4 & 2 & 4 \\ b & 4 & 0 & 3 & 5 \\ c & 2 & 3 & 0 & 2 \\ d & 4 & 5 & 2 & 0 \end{array}$$

How can we find S_2 from S_1? In fact, S_1 MIN $S_1 = S_2$.

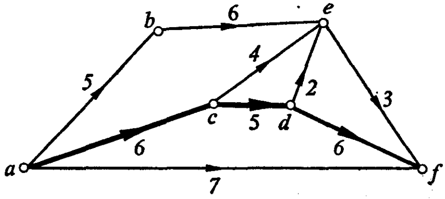
Is the operation MIN associative? Can we do the operation again? What is S_1 MIN S_1 MIN S_1?

This gives us the matrix of shortest routes when we are allowed *one-*, *two-* or *three-*link routes. Notice that, as usually happens, the matrix of shortest routes quickly becomes stable, and no further combination of MIN S_1 alters the entries.

This type of problem is met by London Transport when working out the fares between various parts of London where there are a number of routes and changes between them.

Networks and critical path

8.9 Consider the following network where the numbers represent the times taken on various processes. The problem is to find the maximum time taken and hence the critical path. (See J. Q. Thompson,



An Elementary Introduction to Operational Research, Chapter 1; the order of the letters on the following figure has been chosen to comply with Thompson's notation.)

Note that *direction* plays an important part in what follows. The distance (or time) matrix is:

$$M = \begin{array}{c|cccccc} & a & b & c & d & e & f \\ \hline a & 0 & 5 & 6 & - & - & 7 \\ b & - & 0 & - & - & 6 & - \\ c & - & - & 0 & 5 & 4 & - \\ d & - & - & - & 0 & 2 & 6 \\ e & - & - & - & - & 0 & 3 \\ f & - & - & - & - & - & 0 \end{array}$$

or, in more usual terminology,

$$M = \begin{pmatrix} 0 & 5 & 6 & - & - & 7 \\ - & 0 & - & - & 6 & - \\ - & - & 0 & 5 & 4 & - \\ - & - & - & 0 & 2 & 6 \\ - & - & - & - & 0 & 3 \\ - & - & - & - & - & 0 \end{pmatrix}$$

We now find M MAX M by a process similar to the previous one. We proceed as if we were going to *multiply* the matrices but instead we take *first row* and *first column* etc. *add* corresponding entries and write down the *largest* term. As we are finding the MAX, the – would be replaced by $-\infty$ if a computer were used (in practice -10^6 say).

8.10 For instance, consider *first row* and *fifth column*, i.e.

$$(0 \quad 5 \quad 6 \quad - \quad - \quad 7) \begin{pmatrix} - \\ 6 \\ 4 \\ 2 \\ 0 \\ - \end{pmatrix}$$

Our entry is now the maximum of

$$\{0 + -, 5 + 6, 6 + 4, - + 2, - + 0, - + 7\}$$

i.e. of {–, 11, 10, –, –, –} which is 11.

This, in fact, tells us that the longest *two-link* journey from *a* to *e* is 11. (In fact it tells us the longest *one-* or *two-link* journey is 11.)

Under this rule of combination

$$M \text{ MAX } M = \begin{pmatrix} 0 & 5 & 6 & 11 & 11 & 7 \\ - & 0 & - & - & 6 & 9 \\ - & - & 0 & 5 & 7 & 11 \\ - & - & - & 0 & 2 & 5 \\ - & - & - & - & 0 & 3 \\ - & - & - & - & - & 0 \end{pmatrix}$$

This matrix gives us *all* the longest *one-* or *two-link* journeys. We can continue this process to give us the longest *one-*, *two-* or *three-*link journeys.

M MAX (M MAX M) is

$$\begin{pmatrix} 0 & 5 & 6 & - & - & 7 \\ - & 0 & - & - & 6 & - \\ - & - & 0 & 5 & 4 & - \\ - & - & - & 0 & 2 & 6 \\ - & - & - & - & 0 & 3 \\ - & - & - & - & - & 0 \end{pmatrix} \text{ MAX } \begin{pmatrix} 0 & 5 & 6 & 11 & 11 & 7 \\ - & 0 & - & - & 6 & 9 \\ - & - & 0 & 5 & 7 & 11 \\ - & - & - & 0 & 2 & 5 \\ - & - & - & - & 0 & 3 \\ - & - & - & - & - & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 5 & 6 & 11 & 13 & 17 \\ - & 0 & - & - & 6 & 9 \\ - & - & 0 & 5 & 7 & 11 \\ - & - & - & 0 & 2 & 5 \\ - & - & - & - & 0 & 3 \\ - & - & - & - & - & 0 \end{pmatrix}$$

Is this the same as (M MAX M) MAX M?

8.11 We can repeat this procedure as many times as we like. Remember that we are interested in the *maximum* time from *a* to *d*.

$$M \text{ MAX } (M \text{ MAX } (M \text{ MAX } M)) = \begin{pmatrix} 0 & 5 & 6 & 11 & 13 & \textcircled{17} \\ - & 0 & - & - & 6 & 9 \\ - & - & 0 & 5 & 7 & 11 \\ - & - & - & 0 & 2 & 5 \\ - & - & - & - & 0 & 3 \\ - & - & - & - & - & 0 \end{pmatrix}$$

$$= M \text{ MAX } (M \text{ MAX } M)$$

This means that the process has now stabilized and that the longest time was given by the *three*-link journey, as the *four*-link journeys did not change the maximum times.

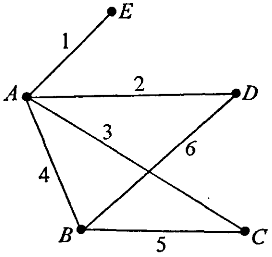
The maximum time *a* to *f* is therefore 17* on a three link journey.

* Draw a flow diagram of this procedure.

It is possible to discover which links are involved by considering how this 17 was arrived at. In this case the answer is obvious from the diagram and our Critical Path is from *a* to *c* to *d* to *f*.

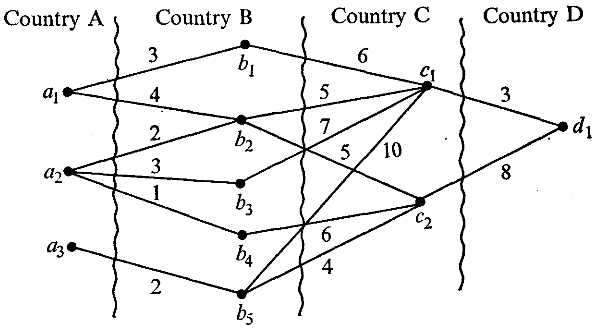
Exercise 8.1

1. Basically a network consists of a number of NODES, joined by a number of LINKS. Here is a typical network. The links are labelled 1, 2, 3 . . . 6. These numbers have nothing to do with lengths, times etc. . . .



It is possible to represent this network in a number of different ways, i.e. either as a *node/node* matrix, a *link/link* matrix or a *node/link* matrix. Write down the matrices in this case and investigate their properties. For example consider the meaning of a *node/node* matrix multiplied by itself; a *node/link* matrix multiplied by (*a*) itself, (*b*) its transpose matrix. Could you draw the network given any of the matrices?

2.

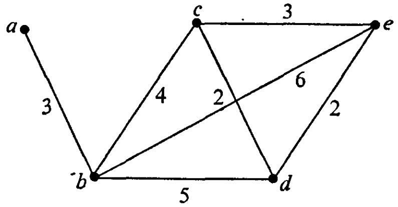


(*a*) How many ways are there from the towns a_1, a_2, a_3 to d_1? How many from $b_1, \ldots, b_5$ to d_1?

(*b*) Suppose the numbers marked represent the distances between the towns. What is the shortest route from a_1 to d_1? From a_2 to d_1? Make up the shortest route matrix.

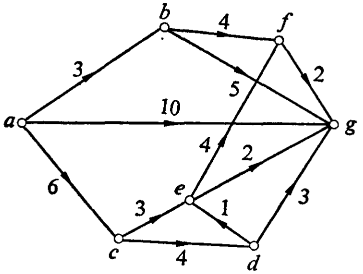
(*c*) Suppose the numbers now represent telephone links (cf. the oil pipe/drain problem earlier). Draw up the maximum link matrix for towns in country A to town d_1 in country *D*.

3. Consider the following network.



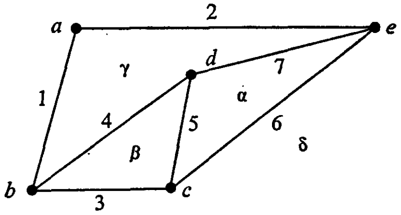
The numbers represent lengths. Write down the one link distance matrix and find out when this matrix stabilizes under the law MIN. Hence give the shortest distance matrix.

4. The following network represents a process in which the times for the various jobs are inserted. Find the critical path from *a* to *g* and the maximum time taken.



5. Look back to the network in **8.2**. In this type of representation can we give any meaning to ADDITION of matrices? Consider the other representations carefully and see if there are addition rules which would mean something in terms of the actual networks.

6. Consider the following network.



This has nodes at *a*, *b*, *c*, *d* and *e*; links 1, 2, 3, 4, 5, 6, 7 and regions α, β, γ, δ.

We can represent this network using a number of different matrices.

(*a*) A node/link matrix.

$$A = \begin{array}{c|ccccccc} & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \hline a & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ b & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ c & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ d & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ e & 0 & 1 & 0 & 0 & 0 & 1 & 1 \end{array}$$

(*b*) A link-region matrix

$$B = \begin{array}{c|cccc} & \alpha & \beta & \gamma & \delta \\ \hline 1 & 0 & 0 & 1 & 1 \\ 2 & 0 & 0 & 1 & 1 \\ 3 & 0 & 1 & 0 & 1 \\ 4 & 0 & 1 & 1 & 0 \\ 5 & 1 & 1 & 0 & 0 \\ 6 & 1 & 0 & 0 & 1 \\ 7 & 1 & 0 & 1 & 0 \end{array}$$

These entries, show for instance, that region α is bounded by links 5, 6 and 7.

(*c*) A node/region matrix

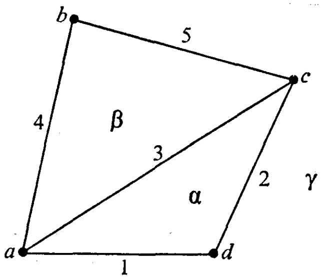
$$C = \begin{array}{c|cccc} & \alpha & \beta & \gamma & \delta \\ \hline a & 0 & 0 & 1 & 1 \\ b & 0 & 1 & 1 & 1 \\ c & 1 & 1 & 0 & 1 \\ d & 1 & 1 & 1 & 0 \\ e & 1 & 0 & 1 & 1 \end{array}$$

Work out the values of

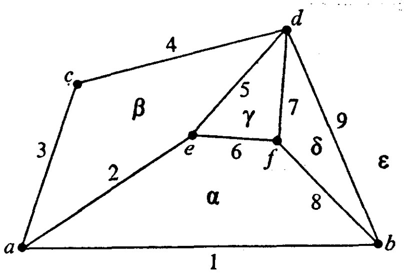
$$AA',\ A'A,\ BB',\ B'B,\ CC' \text{ and } C'C$$

Interpret your answer in each case.

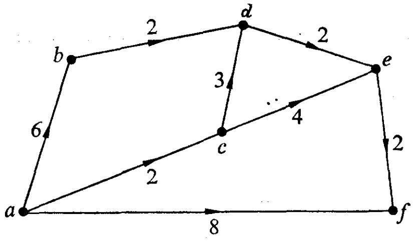
7. Calculate A, B and C for the network below.



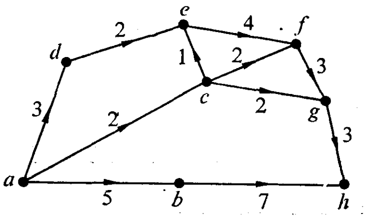
8. As question 7 for the network.



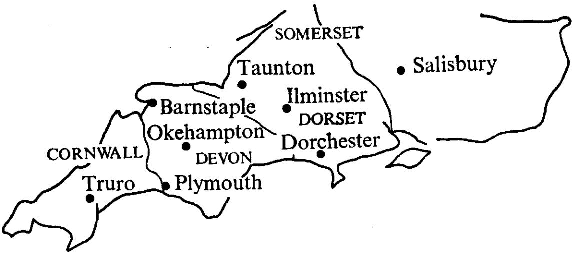
9. Find the critical path for the following network, going from a to f.



10. Find the critical path for the following network going from a to h.



11.



Above is an approximate map of South-West England. Use matrices to find the shortest route from Truro to Salisbury passing through each of the areas shown. The mileages are

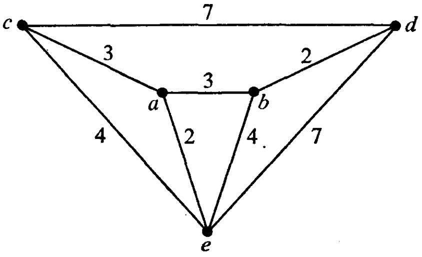
	Barn.	Ok.	Ply.
Truro	87	66	55

	Tau.	Ilm.	Dorch.
Barn.	50	66	86
Ok.	55	57	76
Ply.	74	76	94

	Sal.
Tau.	64
Ilm.	43
Dorch.	40

12. Project. Find out what you can about *dual* configurations. Write down their matrices and see if you can find how (or if) they are connected (a book on projective geometry might help). Look up Bow's notation in a mechanics book (e.g. Humphrey, *Statics*) and find out how it works.

13. The network represents a number of shops belonging to a firm in a small area. The problem is to decide which of the two shops *a* or *b* should be used as a central store. We really want to know if the total of the shortest distances from *a* to *b*, *a* to *c*, *a* to *d* etc. is less than the total of the shortest distances from *b* to *a*, *b* to *c*, *b* to *d* etc.

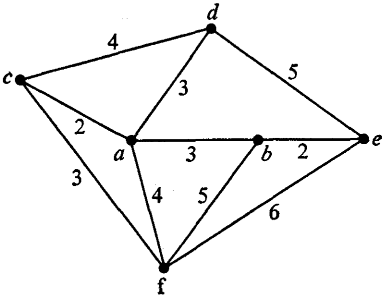


The one link distance matrix is

$$S = \begin{array}{c|ccccc} & a & b & c & d & e \\ \hline a & 0 & 3 & 3 & - & 2 \\ b & 3 & 0 & - & 2 & 4 \\ c & 3 & - & 0 & 7 & 4 \\ d & - & 2 & 7 & 0 & 7 \\ e & 2 & 4 & 4 & 7 & 0 \end{array}$$

Find S MIN S, and S MIN (S MIN S). Deduce that a is the most central store required and that the total distance is 2 miles less than from b.

14. Which of a or b is the most central in the network below? (See No. 13.)



9

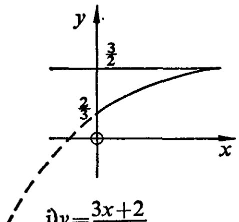
Further Situations in which Matrices are Used

9.1 This chapter contains information about other situations in which matrices are used. The list is not comprehensive and most of the topics are not studied in depth. The topics considered are:

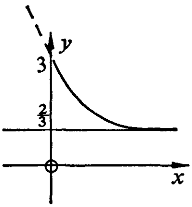
- I An example from the theory of numbers.
- II Recurrence relations.
- III Difference equations.
- IV Solution of differential equations.
- V Four-terminal networks.
- VI Affine transformations.
- VII Matrices and the theory of games.
- VIII Knots.
- IX Matrices with complex entries.

I AN EXAMPLE FROM THE THEORY OF NUMBERS*

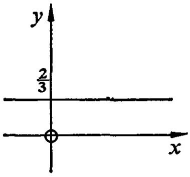
9.2 If $y = \dfrac{ax+b}{cx+d}$ with $a, b, c, d, x > 0$, then y lies between $\dfrac{a}{c}$ and $\dfrac{b}{d}$.



i) $y = \dfrac{3x+2}{2x+3}$



ii) $y = \dfrac{2x+3}{3x+1}$



iii) $y = \dfrac{2x+2}{3x+3} = 2/3$

If

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

---

* cf. *Matrices 2* by G. Matthews, p. 64.

then

(i) $|A| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} > 0$

(ii) $|A| < 0$, and

(iii) $|A| = 0$.

Assuming that $|A| \neq 0$, we can write

$$y = \frac{ax + b}{cx + d} \quad \text{as} \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \bullet x$$

or

$$y = A \bullet x$$

This defines a new operation $\bullet$

9.3 Suppose now that $x = B \bullet z$, where

$$B = \begin{pmatrix} p & q \\ r & s \end{pmatrix}, \quad |B| \neq 0 \quad \text{and} \quad p, q, r, s, z > 0$$

Then

$$x = \frac{pz + q}{rz + s}$$

$$\Rightarrow A \bullet z = \frac{ax + b}{cx + d} = \frac{a(pz + q) + b(rz + s)}{c(pz + q) + d(rz + s)}$$

$$= \frac{(ap + br)z + aq + bs}{(cp + dr)z + cq + ds}$$

$$= \begin{pmatrix} ap + br & aq + bs \\ cp + dr & cq + ds \end{pmatrix} \bullet z$$

$$= (A \,.\, B) \bullet z.$$

Where $A \,.\, B$ is a 'normal' matrix multiplication. Hence

$$A \bullet (B \bullet z) = (A \,.\, B) \bullet z \quad \text{as} \quad x = B \bullet z.$$

9.4 For example, let

$$t = \frac{\sqrt{13} + 3}{2},$$

then

$$t = 3 + \frac{1}{t}$$

$$\Rightarrow t = \frac{3t + 1}{t}$$

i.e.

$$t = \begin{pmatrix} 3 & 1 \\ 1 & 0 \end{pmatrix} \bullet t = C \bullet t$$

where
$$C = \begin{pmatrix} 3 & 1 \\ 1 & 0 \end{pmatrix}$$

Then
$$\begin{aligned} t &= C \bullet t && \text{I} \\ &= C \bullet (C \bullet t) = C^2 \bullet t && \text{II} \\ &= C^2 \bullet (C \bullet t) = C^3 \bullet t && \text{III} \end{aligned}$$

and so on. In fact
$$t = C^n \bullet t.$$

Now
$$C = \begin{pmatrix} 3 & 1 \\ 1 & 0 \end{pmatrix}, \qquad C^2 = \begin{pmatrix} 10 & 3 \\ 3 & 1 \end{pmatrix}$$
$$C^3 = \begin{pmatrix} 33 & 10 \\ 10 & 3 \end{pmatrix}, \qquad C^4 = \begin{pmatrix} 109 & 33 \\ 33 & 10 \end{pmatrix}$$
$$C^5 = \begin{pmatrix} 360 & 109 \\ 109 & 33 \end{pmatrix}$$

and so on.

From II
$$t = \begin{pmatrix} 10 & 3 \\ 3 & 1 \end{pmatrix} \bullet t \Rightarrow \frac{3}{1} < t < \frac{10}{3}$$

From III
$$\begin{aligned} t = \begin{pmatrix} 33 & 10 \\ 10 & 3 \end{pmatrix} \bullet t &\Rightarrow \frac{33}{10} < t < \frac{10}{3} \\ &\Rightarrow \frac{109}{33} < t < \frac{33}{10} \\ &\Rightarrow \frac{360}{109} < t < \frac{109}{33} \text{ and so on.} \end{aligned}$$

9.5 Our next problem is to evaluate C^n. To do this we must find the eigenvector and corresponding eigenvalues of C.*

We have to solve
$$\begin{pmatrix} 3 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = k \begin{pmatrix} x \\ y \end{pmatrix}$$

i.e.
$$\left.\begin{aligned} 3x + y &= kx \\ x &= ky \end{aligned}\right\} \text{ to find } k$$
$$\Leftrightarrow \left.\begin{aligned} (k-3)x &= y \\ x &= ky \end{aligned}\right\}$$

---

* cf. *Matrices 2*, p. 46ff.

These equations are consistent if and only if

$$\frac{k-3}{1} = \frac{1}{k} \Leftrightarrow k^2 - 3k - 1 = 0$$

$$k_1 = \frac{3+\sqrt{13}}{2}$$

$$k_2 = \frac{3-\sqrt{13}}{2}$$

Taking

$$k_1 = \frac{3+\sqrt{13}}{2}.$$

we have

$$\frac{\sqrt{13}-3}{2}\,.\,x = y$$

or

$$\sqrt{13} - 3)x = 2y$$

so that a corresponding eigenvector would be

$$\begin{pmatrix} 2 \\ \sqrt{13}-3 \end{pmatrix}$$

Similarly for

$$k_2 = \frac{3-\sqrt{13}}{2}$$

we have $(-\sqrt{13} - 3)x = 2y$ and the corresponding eigenvector would be

$$\begin{pmatrix} 2 \\ -\sqrt{13}-3 \end{pmatrix}$$

We now use $CU = UD$, where

$$C = \begin{pmatrix} 3 & 1 \\ 1 & 0 \end{pmatrix}, \qquad U = \begin{pmatrix} 2 & 2 \\ \sqrt{13}-3 & -\sqrt{13}-3 \end{pmatrix}$$

and

$$D = \begin{pmatrix} \frac{3+\sqrt{13}}{2} & 0 \\ 0 & \frac{3-\sqrt{13}}{2} \end{pmatrix}.$$

Then

$$\begin{aligned} C &= UDU^{-1} \\ C^2 &= U\,.\,DU^{-1}\,.\,UDU^{-1} \\ &= UDIDU^{-1} \\ &= UD^2U^{-1} \end{aligned}$$

and so on, so that

$$C^n = UD^nU^{-1}$$

$$= \begin{pmatrix} 2 & 2 \\ \sqrt{13}-3 & -\sqrt{13}-3 \end{pmatrix} \begin{pmatrix} \left(\dfrac{3+\sqrt{13}}{2}\right)^n & 0 \\ 0 & \left(\dfrac{3-\sqrt{13}}{2}\right)^n \end{pmatrix}$$

$$- \frac{1}{4\sqrt{13}} \begin{pmatrix} -\sqrt{13}-3 & -2 \\ 3-\sqrt{13} & 2 \end{pmatrix}$$

i.e.

$$C^n = -\frac{1}{4\sqrt{13}} \begin{pmatrix} 2 & 2 \\ \sqrt{13}-3 & -\sqrt{13}-3 \end{pmatrix} \begin{pmatrix} t^n & 0 \\ 0 & u^n \end{pmatrix} \begin{pmatrix} -\sqrt{13}-3 & -2 \\ 3-\sqrt{13} & 2 \end{pmatrix}$$

where

$$t = \frac{3+\sqrt{13}}{2}$$

and

$$u = \frac{3-\sqrt{13}}{2}$$

This can be simplified further to give

$$C^n = \frac{-1}{4\sqrt{13}} \begin{pmatrix} 2 & 2 \\ -2u & -2t \end{pmatrix} \begin{pmatrix} -2t^{n+1} & -2t^n \\ 2u^{n+1} & 2u^n \end{pmatrix}$$

$$= -\frac{1}{\sqrt{13}} \begin{pmatrix} -\ t^{n+1} + u^{n+1} & -t^n + u^n \\ +ut^{n+1} - u^{n+1}t & ut^n - tu^n \end{pmatrix}$$

$$= \frac{1}{\sqrt{13}} \begin{pmatrix} t^{n+1} - u^{n+1} & t^n - u^n \\ ut(u^n - t^n) & ut(u^{n-1} - t^{n-1}) \end{pmatrix}$$

and $t - u = \sqrt{13}$, $ut = -1$. Show that

$$|C^n| = \pm 1$$

and that

$$(t^{n+1} - u^{n+1})(t^{n-1} - u^{n-1}) - (t^n - u^n)^2 = 13(-1)^n$$

9.6 This method can be used as an approximation process for finding the square root of a number* since for example consider

$$t = \begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix} \bullet t$$

i.e.

$$t = \frac{t+3}{t+1}$$

$$\underline{\Rightarrow t^2 = 3}$$

---

* See also p. 128.

This idea is to guess a value for $\sqrt{3}$, 2 say, which we write as $\begin{pmatrix}2\\1\end{pmatrix}$ and then find

$$\begin{pmatrix}1 & 3\\1 & 1\end{pmatrix}\begin{pmatrix}2\\1\end{pmatrix}=\begin{pmatrix}5\\3\end{pmatrix}$$

and to use $\begin{pmatrix}5\\3\end{pmatrix}$ or $^5/_3$ as a second approximation. $\frac{5}{3}\approx 1.66$.

The process is then continued

$$\begin{pmatrix}1 & 3\\1 & 1\end{pmatrix}\begin{pmatrix}5\\3\end{pmatrix}=\begin{pmatrix}14\\8\end{pmatrix}, \frac{14}{8}\approx 1.750.$$

$$\begin{pmatrix}1 & 3\\1 & 1\end{pmatrix}\begin{pmatrix}14\\8\end{pmatrix}=\begin{pmatrix}38\\22\end{pmatrix}, \frac{38}{22}\approx 1.7273.$$

and so on, i.e.

$$\begin{pmatrix}1 & 3\\1 & 1\end{pmatrix}\begin{pmatrix}x_{n-1}\\y_{n-1}\end{pmatrix}=\begin{pmatrix}x_n\\y_n\end{pmatrix}$$

or

$$\begin{pmatrix}x_n\\y_n\end{pmatrix}=\begin{pmatrix}1 & 3\\1 & 1\end{pmatrix}^n\begin{pmatrix}x_0\\y_0\end{pmatrix}$$

where $\begin{pmatrix}x_0\\y_0\end{pmatrix}$ is the first guess.

A further application: bilinear transformations

9.7 Bilinear transformations are transformations of the type

$$w=\frac{az+b}{cz+d},$$

where w and z are complex numbers and a, b, c, d are such that

$$ad-bc\neq 0$$

Transformations of this type are extremely valuable in electricity theory in the same way as affine transformations.* This type of transformation is the simplest which ensures a 1–1 correspondence of the points in the w plane with points in the z plane.

Consider for example, the transformation

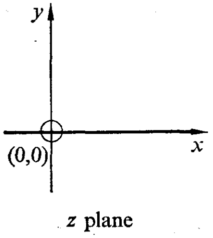
$$w=\frac{2z-1}{z+2}$$

where $w=u+jv$, and $z=x+jy$.

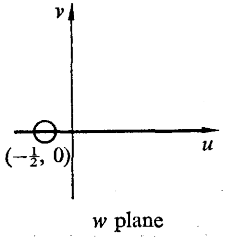
---

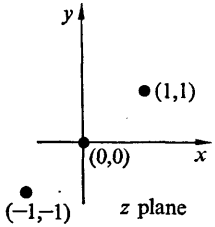
* See **9.34.**

Here are some facts about the transformation

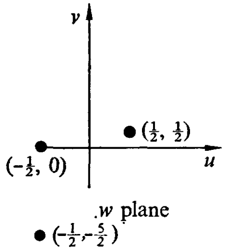


transforms to

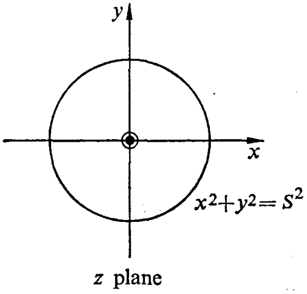




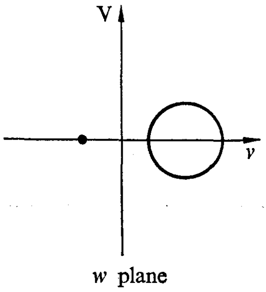
transform to



Points on the line $y = x$ transform to give three non-collinear points. Is this true in general?



transforms to



Investigate this transformation further.

Answer the following question.

A transformation from x to y is given by the relation

$$y = \frac{ax + b}{cx + d}$$

all numbers being real. It is written briefly in the form $y = T(x)$. The 'inverse operation', expressing x in terms of y, is written $x = T^{-1}(y)$.

Prove that T^{-1} is of the same form as T and find the corresponding numbers a', b', c' and d'.

Prove also that, if T_1, given by a_1, b_1, c_1, d_1, and T_2, given by a_2, b_2, c_2, d_2, are two such transformations, then $T_1\{T_2(x)\}$ is also of the same type.

Find the conditions satisfied by a, b, c, d if

$$T\{T(x)\} \equiv x,$$

identically for all values of x. (*CS*)

9.8 Remember that these transformations are, in general

$$w = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \bullet z$$

What type of transformations correspond to $\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$? Show that the inverse transformation, i.e. of $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, is $\dfrac{1}{ad-bc}\begin{pmatrix} +d & -b \\ -c & +a \end{pmatrix}$.

Further information on bilinear transformations can be found in books on the complex variable, e.g. *Functions of a Complex Variable* by Phillips; and in books on electro-magnetism, e.g. *Electro-magnetic Theory* by Ferraro.

These ideas also occur in sequences which satisfy recurrence relations, e.g.

$$u_{n+1} = \frac{au_n + b}{cu_n + d}$$

Is it possible to apply our ideas in this case?

II RECURRENCE RELATIONS*

9.9 We have seen that a recurrence relation of the type

$$u_{r+1} = \frac{au_r + b}{cu_r + d}$$

can be written

$$u_{r+1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \bullet u_r$$

leading to $u_n = A^n \bullet u_0$, where

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

---

* We are indebted to Dr. T. J. Fletcher for the ideas used in this section.

Once again we can find A^n using the ideas of eigenvalues and eigenvectors, and given u_0 we can solve the recurrence relation.

9.10 Occasionally we come across simultaneous difference equations —for example when we are dealing with chemical reactions. These relations can be worked using traditional methods or by the use of matrices.

Examples

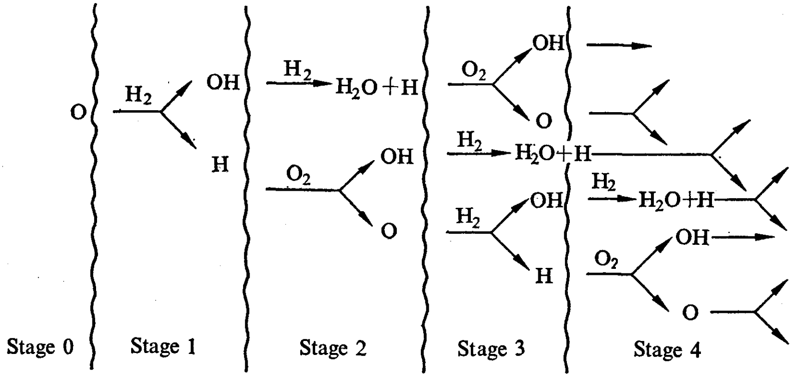
Hydrogen and oxygen react as follows

$$2H_2 + O_2 \rightarrow 2H_2O$$

This is a simplified version of what actually happens. The reaction only occurs because of the presence of O, H and OH. There are three simultaneous reactions:

$$\begin{aligned} O + H_2 &\rightarrow OH + H \\ OH + H_2 &\rightarrow H_2O + H \\ H + O_2 &\rightarrow OH + O \end{aligned}$$

A single O starts a chain reaction as follows:



Suppose that the reactions take place at the same speed and that there is an unlimited supply of O_2 and H_2 molecules. Initially we start the reaction with one oxygen atom and find out how many O, OH and H there are at each successive stage. Let the numbers of these was at stage n be O_n, OH_n and H_n respectively. Then

$$\begin{aligned} O_{n+1} &= H_n \\ OH_{n+1} &= O_n + H_n \\ H_{n+1} &= O_n + OH_n \end{aligned}$$

with the initial conditions

$$O_0 = 1, \quad OH_0 = 0 \quad \text{and} \quad H_0 = 0.$$

We can write these equations in the form

$$\begin{pmatrix} O_{n+1} \\ OH_{n+1} \\ H_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} O_n \\ OH_n \\ H_n \end{pmatrix}$$

or $$Z_{n+1} = AZ_n.$$

Hence $$Z_n = A^n Z_0.$$

Notice that

$$O_{n+1} = 2O_{n-1} + O_{n-2}.$$
$$H_{n+1} = 2H_{n-1} + H_{n-2}$$

and $$OH_{n+1} = 2OH_{n-1} + OH_{n-2}$$

and this leads to

$$(A^3 - 2A - I)Z_0 = O.$$

This is true for any Z_0 and hence

$$A^3 - 2A - I = O.$$

9.11 This is the Cayley-Hamilton theorem—every matrix satisfies its own characteristic equation. We can obtain the characteristic equation as follows:

$$\begin{vmatrix} -\lambda & 0 & 1 \\ 1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{vmatrix} = 0 \quad \text{or} \quad \lambda^3 - 2\lambda - 1 = 0$$

The problem is now the usual one of finding A^n. We do this in the usual way, finding eigenvectors and eigenvalues to use:

$$A = UDU^{-1}, \quad A^n = UD^nU^{-1}$$

where D is a diagonal matrix of eigenvalues and U is the matrix with corresponding eigenvectors as columns. In this case D is

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & \dfrac{-1+\sqrt{5}}{2} & 0 \\ 0 & 0 & \dfrac{-1-\sqrt{5}}{2} \end{pmatrix}$$

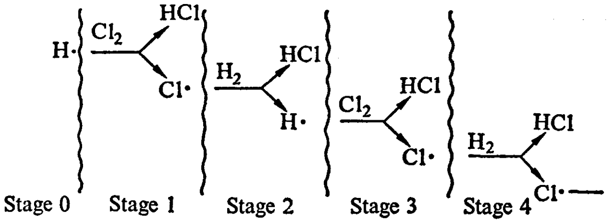
9.12 A simple example of this kind occurs in the reaction between hydrogen and chlorine,

$$H_2 + Cl_2 \rightarrow 2HCl.$$

This is made up of two simultaneous reactions

$$H\cdot + Cl_2 \rightarrow HCl + Cl\cdot$$
$$Cl\cdot + H_2 \rightarrow HCl + H\cdot$$

and can be illustrated as:



This leads to the equations

$$H_n = Cl_{n-1}$$
$$Cl_n = H_{n-1}$$

or

$$\begin{pmatrix} H_n \\ Cl_n \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} H_{n-1} \\ Cl_{n-1} \end{pmatrix}$$

The eigenvalues are $+1$, -1 and hence it is easy to find

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^n;$$

in fact it is either

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

depending whether n is even or odd.

Knowing the initial conditions $H_0 = 1$, $Cl_0 = 0$ we can now solve the problem entirely.

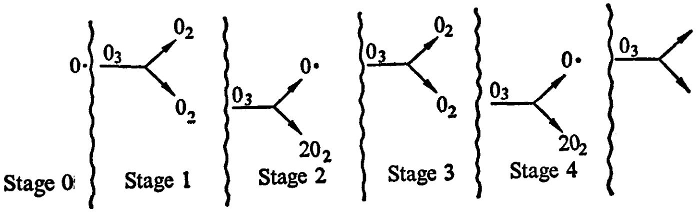
9.13 Solve the two following chain reactions.

1. $$2O_3 \rightarrow 3O_2$$

Chain initiating reaction

$$O_3 \rightarrow O_2 + O\cdot$$

Chain mechanism

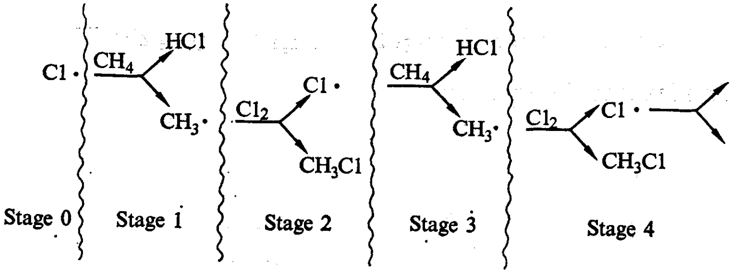


2. $$CH_4 + Cl_2 \rightarrow CH_3Cl + HCl$$

Chain initiating reaction

$$Cl_2 \xrightarrow{\text{Heat or light}} 2Cl\cdot$$

Chain mechanism



Note that these chain reactions are similar in some ways to those which occur in nuclear fission. In the fission of uranium, neutrons are the propagating particles. However, elementary examples are not easy to find.

III DIFFERENCE EQUATIONS

9.14 In more advanced algebra, it is frequently necessary to solve 'difference' equations of the type

$$u_{n+1} - 4u_n + 3u_{n-1} = 0$$

given that $u_0 = 2$, $u_1 = 3$.

At first glance it is not obvious how equations of this sort may be solved using matrices, but in fact by using a simple substitution the method becomes clear. Put

$$v_{n+1} = u_n$$

This gives

$$u_{n+1} = 4u_n - 3v_n$$
$$v_{n+1} = u_n$$

so

$$\begin{pmatrix} u_{n+1} \\ v_{n+1} \end{pmatrix} = \begin{pmatrix} 4 & -3 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} u_n \\ v_n \end{pmatrix}$$

Similarly

$$\begin{pmatrix} u_n \\ v_n \end{pmatrix} = \begin{pmatrix} 4 & -3 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} u_{n-1} \\ v_{n-1} \end{pmatrix}$$

and so on, until

$$\begin{pmatrix} u_2 \\ v_2 \end{pmatrix} = \begin{pmatrix} 4 & -3 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ v_1 \end{pmatrix}$$

giving us

$$\begin{pmatrix} u_{n+1} \\ v_{n+1} \end{pmatrix} = \begin{pmatrix} 4 & -3 \\ 1 & 0 \end{pmatrix}^n \begin{pmatrix} u_1 \\ v_1 \end{pmatrix} \quad \text{where } v_1 = u_0$$

$$= \begin{pmatrix} 4 & -3 \\ 1 & 0 \end{pmatrix}^n \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

9.15 We are then faced with our now familiar problem—that of finding

$$\begin{pmatrix} 4 & -3 \\ 1 & 0 \end{pmatrix}^n$$

The eigenvalues are 3 and 1 so using the method shown in **5.11**, we have

$$\begin{pmatrix} 1^n & 0 \\ 0 & 3^n \end{pmatrix} = a \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$$

$$\Rightarrow a = \frac{3 - 3^n}{2}, \quad b = \frac{3^n - 1}{2}$$

so

$$\begin{pmatrix} 4 & -3 \\ 1 & 0 \end{pmatrix}^n = \frac{3 - 3^n}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{3^n - 1}{2} \begin{pmatrix} 4 & -3 \\ 1 & 0 \end{pmatrix}$$

This gives us

$$\begin{pmatrix} 4 & -3 \\ 1 & 0 \end{pmatrix}^n = \begin{pmatrix} \dfrac{3^{n+1} - 1}{2} & \dfrac{3 - 3^{n+1}}{2} \\ \dfrac{3^n - 1}{2} & \dfrac{3 - 3^n}{2} \end{pmatrix}$$

and hence

$$u_{n+1} = 3 \left(\frac{3^{n+1} - 1}{2} \right) + 2 \left(\frac{3 - 3^{n+1}}{2} \right)$$

$$= \underline{\underline{\frac{3^{n+1}}{2} + \frac{3}{2}}}$$

9.16 This method can, of course, be extended to difference equations of higher order, e.g.

$$u_{n+1} - 2u_n - 5u_{n-1} + 6u_{n-2} = 0.$$

This time the substitutions will be

$$v_{n+1} = u_n$$

and

$$w_{n+1} = v_n.$$

Our equations now become

$$\begin{aligned} u_{n+1} &= 2u_n + 5v_n - 6w_n \\ v_{n+1} &= u_n \\ w_{n+1} &= v_n \end{aligned}$$

or

$$\begin{pmatrix} u_{n+1} \\ v_{n+1} \\ w_{n+1} \end{pmatrix} = \begin{pmatrix} 2 & 5 & -6 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} u_n \\ v_n \\ w_n \end{pmatrix}$$

Exercise 9.1

Solve the following difference equations using the above method.

1. $u_{n+1} - 5u_n + 6u_{n-1} = 0$ with $u_0 = 1$, $u_1 = 2$.
2. $u_{n+1} - 4u_{n-1} = 0$ with $u_0 = 1$, $u_1 = 2$.
3. $u_{n+1} - 3u_n + 2u_{n-1} = 0$ with $u_0 = 2$, $u_1 = 3$.
4. $u_{n+1} - u_n - 12u_{n-1} = 0$ with $u_0 = 3$, $u_1 = 5$.

Consider *carefully* how you would solve

5. $u_{n+1} - 4u_n + 4u_{n-1} = 0$?
6. $u_{n+1} - u_n + 2u_{n-1} = 0$ with $u_0 = 2$, $u_1 = 4$.

9.17 NOTE. Question 5 points to a major difficulty*—that of trying to find the *n*th power of a matrix with equal eigenvalues. These arise in cases where the matrix is of the form

$$M = \begin{pmatrix} 2\alpha & -\alpha^2 \\ 1 & 0 \end{pmatrix}$$

where α is a real number.

To find the n*th power of this matrix*

$$\begin{pmatrix} 2\alpha & -\alpha^2 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 2\alpha & -\alpha^2 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 4\alpha^2 - \alpha^2 & -2\alpha^3 \\ 2\alpha & -\alpha^2 \end{pmatrix} = \begin{pmatrix} 3\alpha^2 & -2\alpha^3 \\ 2\alpha & -\alpha^2 \end{pmatrix} = M^2.$$

Similarly

$$M^3 = \begin{pmatrix} 4\alpha^3 & -3\alpha^4 \\ 3\alpha^2 & -2\alpha^3 \end{pmatrix}$$

$$M^4 = \begin{pmatrix} 5\alpha^4 & -4\alpha^5 \\ 4\alpha^3 & -3\alpha^4 \end{pmatrix}$$

and in general

$$M^n = \begin{pmatrix} (n+1)\alpha^n & -n\alpha^{n+1} \\ n\alpha^{n-1} & -(n-1)\alpha^n \end{pmatrix}.$$

---

* But see the theorem on p. 125.

In Question 5, $\alpha = 2$. Suppose $u_1 = \mathrm{A}$, $u_0 = \mathrm{B}$, then

$$\begin{aligned} u_n &= (n+1)2^n\mathrm{A} - n2^{n+1}\mathrm{B} \\ &= 2^n[\mathrm{A}n + \mathrm{A} - 2\mathrm{B}n] \\ &= 2^n[\mathrm{C}n + \mathrm{D}] \end{aligned}$$

where $\mathrm{A} - 2\mathrm{B} = \mathrm{C}$, $\mathrm{D} = \mathrm{A}$.

This is the usual form of the general solution in such cases.

IV SOLUTION OF DIFFERENTIAL EQUATIONS

9.18 Consider for example, the equation

$$\frac{d_2x}{dt^2} - 5\frac{dx}{dt} + 6x = 0.$$

The first problem is to put this in matrix form. This is done by substituting

$$\frac{dx}{dt} = y.$$

Then

$$\frac{dy}{dt} = 5y - 6x$$

so

$$\begin{pmatrix} \dfrac{dx}{dt} \\ \dfrac{dy}{dt} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -6 & 5 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

which we write as

$$\frac{d}{dt}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -6 & 5 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

Consider the particular case where

$$\frac{d}{dt}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -6 & 5 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

and suppose that when $t = 0$,

$$x = \alpha, \quad \frac{dx}{dt} = \beta.$$

Now

$$\int \frac{d}{dt}\begin{pmatrix} x \\ y \end{pmatrix} dt = \int \begin{pmatrix} 0 & 1 \\ -6 & 5 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} dt$$

hence

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \int_0^t \begin{pmatrix} 0 & 1 \\ -6 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} d\phi$$

where ϕ is a variable of integration.

We now substitute for $\begin{pmatrix} x \\ y \end{pmatrix}$ i.e.

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \int_0^t \begin{pmatrix} 0 & 1 \\ -6 & 5 \end{pmatrix} \left\{ \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \int_0^t \begin{pmatrix} 0 & 1 \\ -6 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} d\phi \right\} d\phi$$

and so on—substituting again and again . . . for $\begin{pmatrix} x \\ y \end{pmatrix}$

$$\begin{pmatrix} x \\ y \end{pmatrix} = \left[I + \int_0^t \begin{pmatrix} 0 & 1 \\ -6 & 5 \end{pmatrix} d\phi + \int_0^t \begin{pmatrix} 0 & 1 \\ -6 & 5 \end{pmatrix} \left\{ \int_0^t \begin{pmatrix} 0 & 1 \\ -6 & 5 \end{pmatrix} d\phi \right\} d\phi + \dots \right] \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

Now in this case our integration is straightforward since the elements of $\begin{pmatrix} 0 & 1 \\ -6 & 5 \end{pmatrix}$ are all constants. Hence

$$\begin{pmatrix} x \\ y \end{pmatrix} = \left[I + t \begin{pmatrix} 0 & 1 \\ -6 & 5 \end{pmatrix} + \frac{t^2}{2!} \begin{pmatrix} 0 & 1 \\ -6 & 5 \end{pmatrix}^2 + \frac{t^3}{3!} \begin{pmatrix} 0 & 1 \\ -6 & 5 \end{pmatrix}^3 + \dots \right] \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

so that

$$\begin{pmatrix} x \\ y \end{pmatrix} = e^{\begin{pmatrix} 0 & 1 \\ -6 & 5 \end{pmatrix} t} \cdot \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

Now by Cayley-Hamilton theorem

$$e^{\begin{pmatrix} 0 & 1 \\ -6 & 5 \end{pmatrix} t} = b_0 I + b_1 \begin{pmatrix} 0 & 1 \\ -6 & 5 \end{pmatrix}$$

and as the eigenvalues are given by

$$\Rightarrow \lambda = 3, \quad \lambda = 2 \text{ since} \begin{vmatrix} -\lambda & 1 \\ -6 & 5 - \lambda \end{vmatrix} = 0$$

we have

$$e^{3t} = b_0 + 3b_1$$

and

$$e^{2t} = b_0 + 2b_1$$

Giving us the solution

$$\begin{pmatrix} x \\ \dfrac{dx}{dt} \end{pmatrix} = \begin{pmatrix} 3e^{2t} - 2e^{3t} & e^{3t} - e^{2t} \\ -6e^{3t} + 6e^{2t} & 3e^{3t} - 2e^{2t} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

9.19 The importance of using matrices to solve differential equations is that in actual problems you usually have *n* equations to solve, or at least an equation of the type

$$\frac{d_n x}{dt^n} + a_1 \frac{d_{n-1}x}{dt^{n-1}} + a_2 \frac{d_{n-2}x}{dt^{n-2}} + \ldots + a_n x = 0.$$

By substitution, as above, this can be reduced to a set of equations such as

$$\begin{pmatrix} \dfrac{dx_1}{dt} \\ \dfrac{dx_2}{dt} \\ \cdot \\ \cdot \\ \cdot \\ \dfrac{dx_n}{dt} \end{pmatrix} = \begin{pmatrix} u_{11} & u_{12} \ldots & u_{1n} \\ u_{21} & u_{22} \ldots & u_{2n} \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ u_{n1} & u_{n2} & u_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix}$$

or

$$\frac{d[\mathbf{x}]}{dt} = M \,.\, [\mathbf{x}].$$

In general, the elements of *M* are functions of *t*, i.e.

$$\frac{d[\mathbf{x}]}{dt} = M(t) \,.\, [\mathbf{x}].$$

Suppose that when $t = 0$,

$$[\mathbf{x}] = [\mathbf{x}_0]$$

$$\int \frac{d[\mathbf{x}]}{dt}\, dt = \int M(t)[\mathbf{x}]\, dt$$

$$[\mathbf{x}] = [\mathbf{x}_0] + \int_0^t M(\phi)[\mathbf{x}]\, d\phi$$

where ϕ is a variable of integration.

9.20 Let us call

$$N = \int_0^t \ldots d\phi$$

i.e. N is an integral operator. Now

$$[\mathbf{x}] = [\mathbf{x}_0] + N(M)[\mathbf{x}].$$

We can substitute for $[\mathbf{x}]$ in this equation to give

$$[\mathbf{x}] = [\mathbf{x}_0] + N(M)[\mathbf{x}_0 + NM[\mathbf{x}]].$$

By continued substitution we have

$$[\mathbf{x}] = (I + N(M) + N(M)N(M) + \ldots)[\mathbf{x}_0]$$

where

$$N(M) = \int_0^t (M)\,d\phi, \quad N(M)N(M) = \int_0^t (M) \int_0^t (M)\,d\phi\,d\phi.$$

If the series

$$I + N(M) + N(M)N(M) + \ldots$$

is convergent let us call the sum of this series a new matrix $P(M)$, i.e.

$$P(M) = I + N(M) + N(M)N(M) + \ldots$$

So $\qquad [\mathbf{x}] = P(M)[\mathbf{x}_0].$

If the elements of M are constants, we can integrate to give

$$\begin{aligned} P(M) &= I + \frac{(M)t}{1!} + \frac{(M^2)t^2}{2!} + \frac{(M^3)t^3}{3!} + \ldots \\ &= e^{(M)t}. \end{aligned}$$

Hence

$$e^{(M)t} = a_0 I + a_1 M + a_2 M^2 + \ldots$$

where $a_0, a_1, a_2 \ldots$ are constant.

If M is a 2×2 matrix, by the Cayley-Hamilton theorem we have

$$e^{(M)t} = b_0 I + b_1 M.$$

In the example, the eigenvalues are 3 and 2, so by substitution

$$\left.\begin{aligned} e^{3t} &= b_0 + 3b_1 \\ e^{2t} &= b_0 + 2b_1 \end{aligned}\right\}^{*}$$

and

---

* $e^{Mt} = I + Mt + \frac{M^2t^2}{2!} + \ldots$ Now $M = UDU^{-1}$ where $D = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$ and U can be found.

⇒ By substitution

$$e^{Mt} = U\left[I + Dt + \frac{D^2t^2}{2!} + \ldots\right]U^{-1} = U\begin{pmatrix} e^{3t} & 0 \\ 0 & e^{2t} \end{pmatrix}U^{-1} = Ub_0U^{-1} + b_1UDU^{-1}$$

hence $$\begin{pmatrix} e^{3t} & 0 \\ 0 & e^{2t} \end{pmatrix} = b_0\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b_1\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}.$$

Hence

$$\begin{aligned} b_0 &= 3e^{2t} - 2e^{3t} \\ b_1 &= e^{3t} - e^{2t} \\ \Rightarrow e^{(M)t} &= b_0 I + b_1 (M) \end{aligned}$$

and

$$\begin{pmatrix} x \\ \dfrac{dx}{dt} \end{pmatrix} = \begin{pmatrix} 3e^{2t} - 2e^{3t} & e^{3t} - e^{2t} \\ -6e^{3t} + 6e^{2t} & 8e^{3t} - 7e^{2t} \end{pmatrix} \begin{pmatrix} x_0 \\ \left(\dfrac{dx}{dt}\right)_0 \end{pmatrix}$$

which gives $x = \mathrm{A}e^{2t} + \mathrm{B}e^{3t}$ as usual.

Note on differentiation and integration of matrices

9.21 Consider the matrix

$$L(x) = \begin{pmatrix} f(x) & g(x) \\ h(x) & i(x) \end{pmatrix}$$

then the differential of $L(x)$ will be, by definition,

$$\operatorname*{Lim}_{\delta x \to 0} \frac{L(x + \delta x) - L(x)}{\delta x}$$

$$= \operatorname*{Lim}_{\delta x \to 0} \begin{pmatrix} \dfrac{f(x + \delta x) - f(x)}{\delta x} & \dfrac{g(x + \delta x) - g(x)}{\delta x} \\ \dfrac{h(x + \delta x) - h(x)}{\delta x} & \dfrac{i(x + \delta x) - i(x)}{\delta x} \end{pmatrix}$$

$= \begin{pmatrix} f'(x) & g'(x) \\ h'(x) & i'(x) \end{pmatrix}$, provided $f(x)$, $g(x)$, $h(x)$, $i(x)$ are differentiable functions of x.

The differentiation of a product can likewise be carried out to give

$$\frac{d}{dx} L_1(x)L_2(x) = L_1(x)\frac{d}{dx} L_2(x) + \frac{d}{dx} L_1(x) \,.\, L_2(x).$$

Check this by taking as an example

$$\begin{pmatrix} 3 & 2x \\ x^3 & x^2 \end{pmatrix} \begin{pmatrix} \cos x & 0 \\ \sin x & x \end{pmatrix}$$

and differentiating.

Integration (seen as the opposite process to differentiation) will give

$$\int L(x)\, dx = \begin{pmatrix} \displaystyle\int_0^x f(t)\, dt & \displaystyle\int_0^x g(t)\, dt \\ \displaystyle\int_0^x h(t)\, dt & \displaystyle\int_0^x i(t)\, dt \end{pmatrix}$$

where t is a variable of integration.

9.22 The problems* which occur are those when the eigenvalues of M are not distinct, e.g. in the case such as

$$\frac{d_2x}{dt^2} - 4\frac{dx}{dt} + 4x = 0$$

i.e.

$$\begin{pmatrix} \dfrac{dx}{dt} \\ \dfrac{dy}{dt} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -4 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \quad \text{where} \quad y = \frac{dx}{dt}.$$

The theory discussed is valid but it is impossible from the method given to find $e^{Mt}(= b_0I + b_1M)$ since we only manage to get one equation twice. This means that we shall have to use another method for finding e^{Mt}.

Suppose we are dealing with 2 × 2 matrices and the equal eigenvalues are λ, then if our matrix is $M\left\{\begin{pmatrix} 0 & 1 \\ -4 & 4 \end{pmatrix} \text{ in this case}\right\}$, we have

$$M^n = n\lambda^{n-1}M - (n-1)\lambda^nI$$

$$\text{Now } e^M = I + M + \frac{M^2}{2!} + \frac{M^3}{3!} + \frac{M^4}{4!} + \dots$$

$$= I + M + \frac{1}{2!}[2\lambda M - \lambda^2I] + \frac{1}{3!}[3\lambda^2M - 2\lambda^3I]$$

$$+ \frac{1}{4!}[4\lambda^3M - 3\lambda^4I] + \dots$$

$$= I\left[1 - \frac{\lambda^2}{2!} - \frac{2\lambda^3}{3!} - \frac{3\lambda^4}{4!} - \dots\right]$$

$$+ M\left[1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \dots\right]$$

$$= I[e^\lambda(1-\lambda)] + Me^\lambda.$$

[The first series has a general term of

$$\frac{-(n-1)\lambda^n}{n!} = -\left[\frac{1}{(n-1)!} - \frac{1}{n!}\right]\lambda^n$$

This gives the series as $1 - [\lambda e^\lambda - (e^\lambda - 1)]$].

---

* But see the theorem on p. 125.

Hence if t is a constant

$$e^{Mt} = I[e^{\lambda t}(1 - \lambda t)] + Me^{\lambda t}.$$

We have now solved the problem if we know the initial conditions, since if at $t = 0$, $x = \text{A}$, $\frac{dx}{dt} = \text{B}$. Then

$$\begin{pmatrix} x \\ y \end{pmatrix} = e^{Mt}\begin{pmatrix} \text{A} \\ \text{B} \end{pmatrix}$$

and if

$$M = \begin{pmatrix} 0 & 1 \\ -4 & 4 \end{pmatrix}, \quad \lambda = 2$$

$$= \begin{pmatrix} e^{2t}(1 - 2t) & e^{2t} \\ -4e^{2t} & 5e^{2t} - 2t.e^{2t} \end{pmatrix} \cdot \begin{pmatrix} \text{A} \\ \text{B} \end{pmatrix}$$

This leads us, as usual to the general solution in the form of

$$x = e^{2t}(\text{C}t + \text{D}).$$

Exercise 9.2

Use matrices to solve the following differential equations, with initially at $t = 0$, $x = \text{A}$, $\frac{dx}{dt} = \text{B}$.

1. $\frac{d_2x}{dt^2} - 3\frac{dx}{dt} + 2x = 0,$
2. $\frac{d_2x}{dt^2} - 6\frac{dx}{dt} + 9x = 0.$
3. $\frac{d_2x}{dt^2} + 4x = 0.$

9.23 We have solved the equations, so far, to give us the *complementary function*, i.e. the solution of an equation where the R.H.S. is zero. We now turn to the problem of the *Particular Integral.*

Suppose our equations are given in the form

$$\text{D}[\mathbf{x}] = M[\mathbf{x}] + [\mathbf{z}]$$

where D means $\frac{d}{dt}$, and M is a matrix [x] and [z] are vectors. We have solved the equation

$$\text{D}[\mathbf{x}] = M[\mathbf{x}]$$

where at time $t = 0$, $[\mathbf{x}] = [\mathbf{x}_0]$, to give

$$[\mathbf{x}] = e^{Mt}[\mathbf{x}_0].$$

Consider now

$$\mathrm{D}[\mathbf{x}] - M[\mathbf{x}] = [\mathbf{z}]$$

where z is a function of t. Now

$$\begin{aligned} \mathrm{D}(e^{-Mt}[\mathbf{x}]) &= -Me^{-Mt}[\mathbf{x}] + e^{-Mt}\,.\,\mathrm{D}[\mathbf{x}] \\ &= e^{-Mt}[\mathrm{D}[\mathbf{x}] - M[\mathbf{x}]] \\ &= e^{-Mt}[\mathbf{z}]. \end{aligned}$$

Integrating we have

$$e^{-Mt}[\mathbf{x}] = \int_{t_0}^{t} e^{-Mt}[\mathbf{z}]\,dt$$

where this vanishes at time t_0, i.e.

$$[\mathbf{x}] = e^{Mt}\int_{t_0}^{t} e^{-Mt}[\mathbf{z}]\,dt.$$

Our complete solution is then

$$\underline{\underline{[\mathbf{x}] = e^{Mt}\left[[\mathbf{x}_0] + \int_{t_0}^{t} e^{-Mt}[\mathbf{z}]\,dt\right]}}$$

9.24 *Example.* To solve the equation

$$\frac{d_2x}{dt^2} - 5\frac{dx}{dt} + 6x = 3t$$

i.e.

$$\frac{d}{dt}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -6 & 5 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ 3t \end{pmatrix} \quad \text{where} \quad y = \frac{dx}{dt}.$$

We have already found the complementary function,* so we only concern ourselves with the particular integral.

This is given by

$$\begin{pmatrix} x \\ y \end{pmatrix} = e^{Mt}\int_{t_0}^{t} e^{-Mt}\begin{pmatrix} 0 \\ 3t \end{pmatrix} dt.$$

Our basic problem then is to find

$$\int_{t_0}^{t} e^{-Mt}\,.\begin{pmatrix} 0 \\ 3t \end{pmatrix} dt.$$

---

* See **9.18**.

This is integrated by parts remembering that when $t = t_0$ the integral is zero.

$$\begin{aligned}\begin{pmatrix}x\\y\end{pmatrix} &= e^{Mt}\left[-\frac{1}{M}e^{-Mt}\begin{pmatrix}0\\3t\end{pmatrix} + \int_{t_0}^{t}\frac{1}{M}e^{-Mt}\cdot\begin{pmatrix}0\\3\end{pmatrix}dt\right].\\
&= e^{Mt}\left[-\frac{1}{M}e^{-Mt}\begin{pmatrix}0\\3t\end{pmatrix} - \frac{1}{M^2}e^{-Mt}\begin{pmatrix}0\\3\end{pmatrix}\right]\\
&= -M^{-1}\begin{pmatrix}0\\3t\end{pmatrix} - M^{-2}\begin{pmatrix}0\\3\end{pmatrix}\\
&= -\left[\frac{1}{6}\begin{pmatrix}5 & -1\\6 & 0\end{pmatrix}\begin{pmatrix}0\\3t\end{pmatrix} + \frac{1}{36}\begin{pmatrix}19 & -5\\30 & -6\end{pmatrix}\begin{pmatrix}0\\3\end{pmatrix}\right]\\
&= -\left[\begin{pmatrix}-\frac{1}{2}t\\0\end{pmatrix} + \begin{pmatrix}-\frac{5}{12}\\-\frac{1}{2}\end{pmatrix}\right]\\
&= \begin{pmatrix}\frac{1}{2}t + \frac{5}{12}\\\frac{1}{2}\end{pmatrix}.\end{aligned}$$

Hence the complete solution is

$$\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}3e^{2t} - 2e^{3t} & e^{3t} - e^{2t}\\-6e^{3t} + 6e^{2t} & 3e^{3t} - 2e^{2t}\end{pmatrix}\begin{pmatrix}x_0\\y_0\end{pmatrix} + \begin{pmatrix}\frac{1}{2}t + \frac{5}{12}\\\frac{1}{2}\end{pmatrix}$$

i.e.

$$x = Ae^{2t} + Be^{3t} + \tfrac{1}{2}t + \tfrac{5}{12}$$

which is the same as we should normally get.

The final problem is to solve an equation such as

$$\frac{d_2x}{dt^2} - 5\frac{dx}{dt} + 6x = e^{5t}.$$

where the right-hand side is of the form of the complementary function and is dealt with in **9.28.**

Exercise 9.3

Find the particular integral in the following cases:

1. $\dfrac{d_2x}{dt^2} - 5\dfrac{dx}{dt} + 6x = 2t^2.$
2. $\dfrac{d_2x}{dt^2} - 5\dfrac{dx}{dt} + 6x = \cos t.$

9.25 We include a different method for solving the equations when they have equal eigenvalues in the matrix. This uses the JORDAN CANONICAL FORM.

This means that any matrix with equal eigenvalues λ can be written as

$$A = B^{-1}\Lambda B$$

where

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \Lambda = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$$

and

$$B = \begin{pmatrix} e & f \\ g & h \end{pmatrix}$$

is to be determined.

Since A has equal eigenvalues

$$(a - \lambda)(d - \lambda) - bc = 0$$
$$\lambda^2 - (a + d)\lambda + ad - bc = 0$$

has equal roots, i.e.

$$\underline{\underline{(a + d)^2 = 4(ad - bc)}} \qquad \text{(i)}$$

Now

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{1}{\Delta}\begin{pmatrix} h & -f \\ -g & e \end{pmatrix}\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}\begin{pmatrix} e & f \\ g & h \end{pmatrix}$$

where $\Delta = eh - fg$

$$= \begin{pmatrix} \lambda + \dfrac{hg}{\Delta} & \dfrac{h^2}{\Delta} \\ -\dfrac{g^2}{\Delta} & \lambda - \dfrac{hg}{\Delta} \end{pmatrix}$$

i.e.

$$a \equiv \lambda + \frac{hg}{\Delta}$$

$$b \equiv \frac{h^2}{\Delta}$$

$$c \equiv -\frac{g^2}{\Delta}$$

$$d \equiv \lambda - \frac{hg}{\Delta}$$

These four values can easily be shown to satisfy (i).

9.26 *Example.* Find the matrix B (this is by no means unique as we shall see) so that:

$$\begin{pmatrix} 2 & 1 \\ -1 & 4 \end{pmatrix} = \frac{1}{\Delta}\begin{pmatrix} h & -f \\ -g & e \end{pmatrix}\begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} e & f \\ g & h \end{pmatrix}$$

i.e. where

$$B = \begin{pmatrix} e & f \\ g & h \end{pmatrix}$$

Our equations become

$$2 = 3 + \frac{hg}{\Delta} \qquad 1 = \frac{h^2}{\Delta}$$

$$-1 = -\frac{g^2}{\Delta} \qquad 4 = 3 - \frac{hg}{\Delta}$$

Effectively we have only three equations

$$\begin{aligned} hg &= -\Delta \\ h^2 &= \Delta \\ g^2 &= \Delta \end{aligned}$$

We can take $\Delta = 1$, then $h = -1$, $g = 1$ are possibilities, since $\Delta = 1$, $1 = eh - fg$, so that $e + f = -1$, and we could consider $e = 4, f = -5$, for instance.

Our solution is

$$\begin{pmatrix} 2 & 1 \\ -1 & 4 \end{pmatrix} = \begin{pmatrix} -1 & 5 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 4 & -5 \\ 1 & -1 \end{pmatrix}$$

This is only one of many possibilities.

9.27 Consider

$$\frac{d}{dt}[\mathbf{x}] = A[\mathbf{x}]$$

where A is a matrix with equal eigenvalues, λ say, then

$$\begin{aligned}
[\mathbf{x}] &= e^{At}[\mathbf{x}_0] \\
&= \left[I + At + \frac{A^2t^2}{2!} + \ldots \right] [\mathbf{x}_0] \text{ where } A = B^{-1}\,\Lambda\,B \\
&= \left[B^{-1} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} B + B^{-1}\,\Lambda t\,B + B^{-1}\frac{(\Lambda t)^2}{2!} B + \ldots \right] [\mathbf{x}_0] \\
&= B^{-1} \left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} t + \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}^2 \frac{t^2}{2!} + \ldots \right] B\,[\mathbf{x}_0] \\
&= B^{-1} \left[\begin{pmatrix} 1 + \lambda t + \dfrac{\lambda^2 t^2}{2!} + \ldots & t + \lambda t^2 + \dfrac{\lambda^2 t^3}{2!} + \ldots \\ 0 & 1 + \lambda t + \dfrac{\lambda^2 t^2}{2!} + \ldots \end{pmatrix} \right] B\,[\mathbf{x}_0] \\
&= B^{-1} \begin{pmatrix} e^{\lambda t} & te^{\lambda t} \\ 0 & e^{\lambda t} \end{pmatrix} B[\mathbf{x}_0]
\end{aligned}$$

and since B^{-1} and B are constants depending on our choice

$$[\mathrm{x}] = \begin{pmatrix} e^{\lambda t} & te^{\lambda t} \\ 0 & e^{\lambda t} \end{pmatrix} \begin{pmatrix} \mathrm{A} \\ \mathrm{B} \end{pmatrix}$$

i.e.

$$x = e^{\lambda t}(\mathrm{A} + \mathrm{B}t)$$

$$\frac{dx}{dt} = \mathrm{B}e^{\lambda t}$$

as usual.

Exercise 9.4

Use this method to solve

1. $\dfrac{d_2x}{dt^2} - 4\dfrac{dx}{dt} + 4x = 0.$
2. $\dfrac{d_2x}{dt^2} + 6\dfrac{dx}{dt} + 9x = 0.$

Problems of the particular integral

9.28 Consider, for example, the equation

$$\frac{d}{dt}[\mathrm{x}] = \begin{pmatrix} 0 & 1 \\ -6 & 5 \end{pmatrix} [\mathrm{x}] + \begin{pmatrix} 0 \\ e^{3t} \end{pmatrix}.$$

We have already found the complementary function* and so we only concern ourselves with the particular integral, i.e.

$$[\mathrm{x}] = e^{Mt} \int_{t_0}^{t} e^{-Mt} \begin{pmatrix} 0 \\ e^{3t} \end{pmatrix} dt \quad \text{where} \quad M = \begin{pmatrix} 0 & 1 \\ -6 & 5 \end{pmatrix}.$$

Now using Cayley-Hamilton theorem as before

$$e^{Mt} = \begin{pmatrix} 3e^{2t} - 2e^{3t} & e^{3t} - e^{2t} \\ -6e^{3t} + 6e^{2t} & 3e^{3t} - 2e^{3t} \end{pmatrix}$$

and

$$e^{-Mt} = \begin{pmatrix} 3e^{-2t} - 2e^{3t} & e^{-3t} - e^{-2t} \\ -6e^{-3t} + 6e^{-2t} & 3e^{-3t} - 2e^{-2t} \end{pmatrix}$$

so

$$e^{-Mt} \begin{pmatrix} 0 \\ e^{3t} \end{pmatrix} = \begin{pmatrix} 1 - e^{t} \\ 3 - 2e^{t} \end{pmatrix}$$

$$\Rightarrow \int_{t_0}^{t} e^{-Mt} \begin{pmatrix} 0 \\ e^{3t} \end{pmatrix} dt = \begin{pmatrix} t - e^{t} \\ 3t - 2e^{t} \end{pmatrix}$$

and the particular integral is

$$\begin{pmatrix} 3e^{2t} - 2e^{3t} & e^{3t} - e^{2t} \\ -6e^{3t} + 6e^{2t} & 3e^{3t} - 2e^{2t} \end{pmatrix} \begin{pmatrix} t - e^{t} \\ 3t - 2e^{t} \end{pmatrix} = \begin{pmatrix} (1 - t)e^{3t} \\ (3t - 2)e^{3t} \end{pmatrix}.$$

---

* See **9.18.**

Hence the complete solution will lead to (as usual),

$$x = \mathrm{A}e^{2t} + (\mathrm{B}t + \mathrm{C})e^{3t}$$

9.29 Consider now an example in which the matrix has equal eigenvalues and the complementary function is of the form required for the particular integral, e.g.

$$\frac{d}{dt}[\mathbf{x}] = \begin{pmatrix} 0 & +1 \\ -9 & 6 \end{pmatrix}[\mathbf{x}] + \begin{pmatrix} 0 \\ e^{3t} \end{pmatrix}$$

$$[\mathbf{x}] = e^{Mt}\int_{t_0}^{t} e^{-Mt}\begin{pmatrix} 0 \\ e^{3t} \end{pmatrix} dt \quad \text{where} \quad M = \begin{pmatrix} 0 & 1 \\ -9 & 6 \end{pmatrix}.$$

The equal eigenvalues are given by $\lambda = 3$. Now

$$e^{-Mt} = I - Mt + \frac{M^2t^2}{2!} - \frac{M^3t^3}{3!} + \dots$$

We could use the Jordan Canonical Form to help us here. Suppose $M = B^{-1}\Lambda B$, where B is to be found as previously,

$$\begin{pmatrix} 0 & 1 \\ -9 & 6 \end{pmatrix} = \begin{pmatrix} 3 + hg & h^2 \\ -g^2 & 3 - hg \end{pmatrix} \quad \text{if} \quad \det B = 1,\ B = \begin{pmatrix} e & f \\ g & h \end{pmatrix}.$$

Hence $\left.\begin{aligned} hg &= -3 \\ h^2 &= 1 \\ g^2 &= 9 \end{aligned}\right\}$ possible solutions being $h = 1$, $g = -3$,

also $eh - fg = 1 \Rightarrow e + 3f = 1$

and we can take $e = -2, f = 1$. Now

$$[\mathbf{x}] = e^{Mt}\int_{t_0}^{t} e^{-Mt}\begin{pmatrix} 0 \\ e^{3t} \end{pmatrix} dt$$

$$= B^{-1}\begin{pmatrix} e^{3t} & te^{3t} \\ 0 & e^{3t} \end{pmatrix} B \int_{t_0}^{t} B^{-1}\begin{pmatrix} e^{-3t} & -te^{-3t} \\ 0 & e^{-3t} \end{pmatrix} B \begin{pmatrix} 0 \\ e^{3t} \end{pmatrix} dt$$

$$= B^{-1}\begin{pmatrix} e^{3t} & te^{3t} \\ 0 & e^{3t} \end{pmatrix} \int_{t_0}^{t} \begin{pmatrix} e^{-3t} & -te^{-3t} \\ 0 & e^{-3t} \end{pmatrix}\begin{pmatrix} -2 & 1 \\ -3 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ e^{3t} \end{pmatrix} dt$$

$$= \begin{pmatrix} 1 & -1 \\ 3 & -2 \end{pmatrix}\begin{pmatrix} e^{3t} & te^{3t} \\ 0 & e^{3t} \end{pmatrix}\int_{t_0}^{t}\begin{pmatrix} 1 - t \\ 1 \end{pmatrix} dt$$

$$= \begin{pmatrix} 1 & -1 \\ 3 & -2 \end{pmatrix}\begin{pmatrix} e^{3t} & te^{3t} \\ 0 & e^{3t} \end{pmatrix}\begin{pmatrix} t - \frac{t^2}{2} \\ t \end{pmatrix}$$

$$= \begin{pmatrix} \frac{t^2}{2}e^{3t} \\ \left[\frac{3t^2}{2} + t\right] e^{3t} \end{pmatrix}$$

This gives us, as we expect $x = \mathrm{A}t^2e^{3t}$.

Exercise 9.5

Use matrices to solve

1. $\frac{d_2x}{dt^2} - 4\frac{dx}{dt} + 4x = 2e^{2t}$.
2. $\frac{d_2x}{dt^2} - 2\frac{dx}{dt} + x = e^{t}$.
3. $\frac{d_2x}{dt^2} - 4\frac{dx}{dt} + 3x = 2e^{t}$.
4. $\frac{d_2x}{dt^2} + 4x = 3 \sin 2t$.

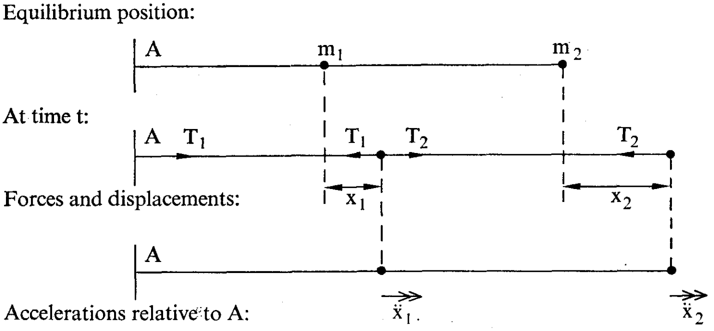
5. Investigate the solutions of third order differential equations using matrices. [Hint the Jordan-Canonical Form gives something like

$$\begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_2 \end{pmatrix}$$

for two equal roots.] What do you think it would be for 3 equal eigenvalues?

The above methods are applicable to all differential equations of the form $\ddot{x} + k_2\dot{x} + k_1x = 0$. Equations of this form occur frequently in mechanical or electrical systems. It is in solving *simultaneous* equations of this form (e.g. in coupled systems) that the power of the method is apparent.

9.30 *Example.* Consider two elastic strings as shown, separating two masses, m_1 and m_2.



If the strings have stiffness factors k_1 and k_2, then by Hooke's Law

$$T_1 = k_1x_1, \quad T_2 = k_2(x_2 - x_1) \tag{1}$$

Using Newton's second law ($P = mf$) for each mass

$$\begin{cases} m_1\ddot{x}_1 = T_2 - T_1 \\ m_2\ddot{x}_2 = -T_2 \end{cases} \tag{2}$$

Using (1)

$$\begin{cases} m_1\ddot{x}_1 = k_2(x_2 - x_1) - k_1x_1 \\ m_2\ddot{x}_2 = -k_2(x_2 - x_1) \end{cases} \tag{3}$$

Hence

$$\begin{cases} \ddot{x}_1 = -\dfrac{(k_1 + k_2)}{m_1}x_1 + \dfrac{k_2}{m_1}x_2 \\ \ddot{x}_2 = +\dfrac{k_2}{m_2}x_1 - \dfrac{k_2}{m_2}x_2 \end{cases} \tag{4}$$

Put $w_1 = \dot{x}_1$ and $w_2 = \dot{x}_2$. Then

$$\begin{cases} \dot{w}_1 = & & \dfrac{-(k_1 + k_2)}{m_1}x_1 + \dfrac{k_2}{m_1}x_2 \\ \dot{x}_1 = & w_1 & \\ \dot{w}_2 = & & \dfrac{k_2}{m_2}x_1 \qquad - \dfrac{k_2}{m_2}x_1 \\ \dot{x}_2 = & & \qquad\qquad w_2 \end{cases} \tag{5}$$

which can be written as

$$\begin{pmatrix} \dot{w}_1 \\ \dot{x}_1 \\ \dot{w}_2 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} 0 & -\dfrac{(k_1 + k_2)}{m_1} & 0 & \dfrac{k_2}{m_1} \\ 1 & 0 & 0 & 0 \\ 0 & \dfrac{k_2}{m_2} & 0 & -\dfrac{k_2}{m_2} \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} w_1 \\ x_1 \\ w_2 \\ x_2 \end{pmatrix} \tag{6}$$

We now find the eigenvalues, λ_1, λ_2, λ_3, λ_4 of this matrix, and compute the corresponding eigenvectors.

Let the matrix A be the matrix whose column vectors are these eigenvectors.

Then the solution to our differential equations can be written

$$\begin{pmatrix} w_1 \\ x_1 \\ w_2 \\ x_2 \end{pmatrix} = A \begin{pmatrix} \mathrm{P}e^{\lambda_1 t} \\ \mathrm{Q}e^{\lambda_2 t} \\ \mathrm{R}e^{\lambda_3 t} \\ \mathrm{S}e^{\lambda_4 t} \end{pmatrix} \tag{7}$$

where P, Q, R, S are constants.

The solutions are of the form

$$x_1 = p_1e^{\lambda_1 t} + q_1e^{\lambda_2 t} + r_1e^{\lambda_3 t} + s_1e^{\lambda_4 t}$$

$$x_2 = p_2e^{\lambda_1 t} + q_2e^{\lambda_2 t} + r_2e^{\lambda_3 t} + s_2e^{\lambda_4 t}$$

Let us suppose $m_1 = m_2 = 1$, $k_1 = 1$ and $k_2 = 2$.

Equation (6) can then be written

$$\begin{pmatrix} \dot{w}_1 \\ \dot{x}_1 \\ \dot{w}_2 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} 0 & -3 & 0 & 2 \\ 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & -2 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} w_1 \\ x_1 \\ w_2 \\ x_2 \end{pmatrix}$$

The eigenvalues of the matrix are then given by

$$\begin{vmatrix} -\lambda & -3 & 0 & 2 \\ 1 & -\lambda & 0 & 0 \\ 0 & 2 & -\lambda & -2 \\ 0 & 0 & 1 & -\lambda \end{vmatrix} = 0$$

$$\Rightarrow \lambda^4 + 5\lambda^2 + 2 = 0$$

$$\Rightarrow \lambda^2 = \frac{-5 \pm \sqrt{25-8}}{2} \qquad = \frac{-5 \pm \sqrt{17}}{2}$$

Hence the eigenvectors are

$$\lambda_1 = j\sqrt{\frac{5+\sqrt{17}}{2}} \qquad \lambda_2 = -j\sqrt{\frac{5+\sqrt{17}}{2}}$$

$$\lambda_3 = j\sqrt{\frac{5-\sqrt{17}}{2}} \qquad \lambda_4 = -j\sqrt{\frac{5-\sqrt{17}}{2}}$$

Let us call these

$$\left.\begin{aligned} \lambda_1 &= mj, \ \lambda_2 = -mj \\ \lambda_3 &= nj, \ \lambda_4 = -nj \end{aligned}\right\}$$

where

$$m = \sqrt{\frac{5+\sqrt{17}}{2}}, \quad n = \sqrt{\frac{5-\sqrt{17}}{2}}$$

Then the required solution is

$$\begin{aligned} x_1 &= p_1e^{jmt} + q_1e^{-jmt} + r_1e^{jnt} + s_1e^{-jnt} \\ x_2 &= p_2e^{jmt} + q_1e^{-jmt} + r_2e^{jnt} + s_2e^{-jnt} \end{aligned}$$

which we usually interpret as

$$\begin{aligned} x_1 &= A_1 \cos mt + B_1 \sin mt + C_1 \cos nt + D_1 \sin nt \\ x_2 &= A_2 \cos mt + B_2 \sin mt + C_2 \cos nt + D_2 \sin nt \end{aligned}$$

where A_1, B_1, C_1, D_1, etc., are arbitrary constants.

Exercise 9.6

1. Solve the preceding problem when

$$m_1 = m_2 = 1, \quad k_1 = 2, \quad k_2 = 1$$

2. In the worked example show that, if each body is damped by a force equal to c times its velocity, then the differential equations describing its motion can be written as

$$\begin{pmatrix} \dot{w}_1 \\ \dot{x}_1 \\ \dot{w}_2 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} -\dfrac{c}{m_1} & -\dfrac{k_1+k_2}{m_1} & 0 & \dfrac{k_2}{m_1} \\ 1 & 0 & 0 & 0 \\ 0 & \dfrac{k_2}{m_2} & -\dfrac{c}{m_2} & -\dfrac{k_2}{m_2} \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} w_1 \\ x_1 \\ w_2 \\ x_2 \end{pmatrix}$$

3. The displacement x of a simple harmonic oscillator satisfies the differential equation

$$\frac{d^2x}{dt^2} + \omega^2 x = 0$$

Denoting by M and M' the function $x = a \cos \omega(t - \alpha)$ and $x = a' \cos \omega(t - \alpha')$, show that the functions λM and $M + M'$, defined by

$$x = \lambda a \cos \omega(t - \alpha)$$

and $$x = a \cos \omega(t - \alpha) + a' \cos \omega(t - \alpha')$$

respectively, satisfy the differential equation. Prove that the solutions of the equation form a vector space. What is its dimension?

The scalar product $M \,.\, M'$ is defined as $aa' \cos \omega(\alpha - \alpha')$. Prove that

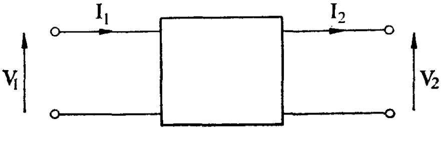
$$(M + M') \,.\, M'' = M \,.\, M'' + M' \,.\, M''$$

Writing $M = m_1 e_1 + m_2 e_2$, where e_1 and e_2 denote the functions $x = \cos \omega t$ and $x = \sin \omega t$ respectively, express $M \,.\, M'$ in terms of the components (m_1, m_2) and (m_1', m_2').

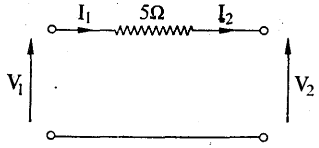
If the time $\bar{t}$ is measured from the instant $t = \theta$, so that $\bar{t} = t - \theta$, and M is written as $\bar{m}_1 \bar{e}_1 + \bar{m}_2 \bar{e}_2$, where $\bar{e}_1$ and $\bar{e}_2$ are the functions $x = \cos \omega \bar{t}$ and $x = \sin \omega \bar{t}$ respectively, find the relation between the components $(\bar{m}_1, \bar{m}_2)$ and (m_1, m_2). What is the nature of the matrix involved? *(CS)*

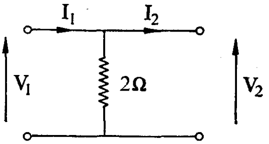
V FOUR-TERMINAL NETWORKS

9.31 A four-terminal network is an electrical system (or black box) with two input and two output terminals, i.e.



Basically all that is required is a knowledge of Ohm's law and common sense, to arrive at the various sets of equations which are satisfied by the voltages and the currents, for example:





$$V_1 = 5I_2 + V_2 \text{ (Ohm's law)} \qquad V_1 = V_2$$
$$I_1 = I_2 \qquad V_2 = 2(I_1 - I_2)$$
$$\text{or } I_1 = \tfrac{1}{2}V_2 + I_2$$

or in matrix notation

$$\begin{pmatrix} V_1 \\ I_1 \end{pmatrix} = \begin{pmatrix} 1 & 5 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} V_2 \\ I_2 \end{pmatrix} \qquad \begin{pmatrix} V_1 \\ I_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} V_2 \\ I_2 \end{pmatrix}$$

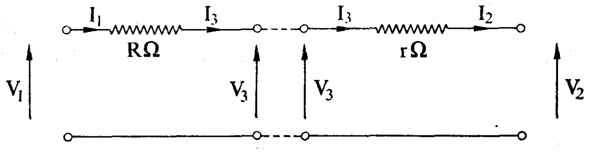
In general for a resistance $R\Omega$ (as on the left above) we have

$$\begin{pmatrix} V_1 \\ I_1 \end{pmatrix} = \begin{pmatrix} 1 & R \\ 0 & 1 \end{pmatrix} \begin{pmatrix} V_2 \\ I_2 \end{pmatrix}$$

and for a resistance as on the right above of SΩ,

$$\begin{pmatrix} V_1 \\ I_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{1}{S} & 1 \end{pmatrix} \begin{pmatrix} V_2 \\ I_2 \end{pmatrix}$$

9.32 We can now apply our ideas to a number of circuits joined up together as, for example:



We now have, for the left hand part

$$\begin{pmatrix} V_1 \\ I_1 \end{pmatrix} = \begin{pmatrix} 1 & R \\ 0 & 1 \end{pmatrix} \begin{pmatrix} V_3 \\ I_3 \end{pmatrix}$$

and for the right hand circuit

$$\begin{pmatrix} V_3 \\ I_3 \end{pmatrix} = \begin{pmatrix} 1 & r \\ 0 & 1 \end{pmatrix} \begin{pmatrix} V_2 \\ I_2 \end{pmatrix}$$

Hence

$$\begin{pmatrix} V_1 \\ I_1 \end{pmatrix} = \begin{pmatrix} 1 & R \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & r \\ 0 & 1 \end{pmatrix}\begin{pmatrix} V_2 \\ I_2 \end{pmatrix}$$

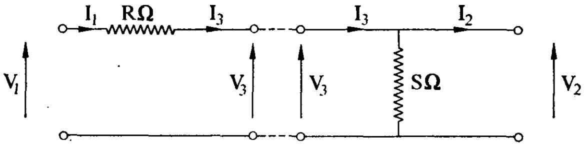
$$= \begin{pmatrix} 1 & R+r \\ 0 & 1 \end{pmatrix}\begin{pmatrix} V_2 \\ I_2 \end{pmatrix}$$

and the rule for adding resistances in series is established. Similarly we can show that for two resistances in parallel the total resistance is given by

$$\frac{1}{R} = \frac{1}{r_1} + \frac{1}{r_2}$$

Prove this for yourself.

We can also combine the two circuits as



Then

$$\begin{pmatrix} V_1 \\ I_1 \end{pmatrix} = \begin{pmatrix} 1 & R \\ 0 & 1 \end{pmatrix}\begin{pmatrix} V_3 \\ I_3 \end{pmatrix}$$

and

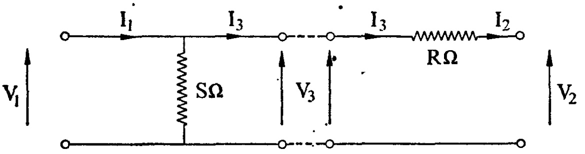
$$\begin{pmatrix} V_3 \\ I_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{1}{S} & 1 \end{pmatrix}\begin{pmatrix} V_2 \\ I_2 \end{pmatrix}$$

so

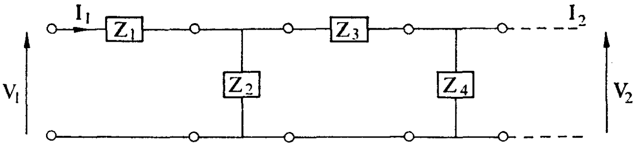
$$\begin{pmatrix} V_1 \\ I_1 \end{pmatrix} = \begin{pmatrix} 1 & R \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ \frac{1}{S} & 1 \end{pmatrix}\begin{pmatrix} V_2 \\ I_2 \end{pmatrix}$$

$$= \begin{pmatrix} 1+\frac{R}{S} & R \\ \frac{1}{S} & 1 \end{pmatrix}\begin{pmatrix} V_2 \\ I_2 \end{pmatrix}$$

9.32 Do we obtain the same results if we combine the circuits in the reverse order? i.e.



In general, if we have a number of circuits combined and the impedance* of each circuit is $Z_1, Z_2, Z_3 \ldots$, i.e.



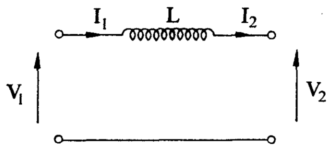
then

$$\begin{pmatrix} V_1 \\ I_1 \end{pmatrix} = \begin{pmatrix} 1 & Z_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \frac{1}{Z_2} & 1 \end{pmatrix} \begin{pmatrix} 1 & Z_3 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \frac{1}{Z_4} & 1 \end{pmatrix} \cdots \begin{pmatrix} V_2 \\ I_2 \end{pmatrix}$$

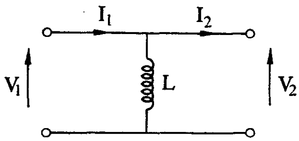
Inductive and capacitive circuits

9.33 In the following circuits, w is the frequency of the applied voltage.

INDUCTIVE



$$\begin{pmatrix} V_1 \\ I_1 \end{pmatrix} = \begin{pmatrix} 1 & wL \\ 0 & 1 \end{pmatrix} \begin{pmatrix} V_2 \\ I_2 \end{pmatrix}$$

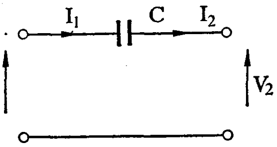


$$\begin{pmatrix} V_1 \\ I_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{1}{wL} & 1 \end{pmatrix} \begin{pmatrix} V_2 \\ I_2 \end{pmatrix}$$

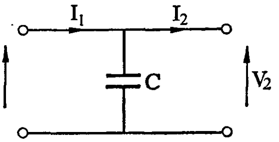
---

* The impedance of a circuit is its resistive effect.

CAPACITIVE



$$\begin{pmatrix} V_1 \\ I_1 \end{pmatrix} = \begin{pmatrix} 1 & \dfrac{1}{wC} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} V_2 \\ I_2 \end{pmatrix}$$



$$\begin{pmatrix} V_1 \\ I_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ wC & 1 \end{pmatrix} \begin{pmatrix} V_2 \\ I_2 \end{pmatrix}$$

Exercise 9.7

1. Investigate the properties of matrices of the type

(*a*) $\begin{pmatrix} 1 & R \\ 0 & 1 \end{pmatrix}$,

(*b*) $\begin{pmatrix} 1 & 0 \\ \dfrac{1}{S} & 1 \end{pmatrix}$

Notice that the determinant of such matrices is $+1$. Do such matrices, under multiplication, form groups. What types of transformations of the plane correspond to these matrices?

2. Many circuits involve cascade circuits, i.e. a large number of similar units joined up. To find the total resistance of a circuit involving n series resistances requires that we can find for example

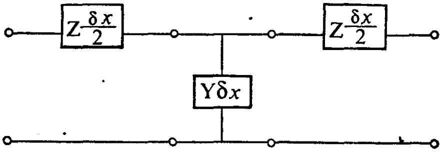
$$\begin{pmatrix} 1 & R \\ 0 & 1 \end{pmatrix}^n$$

How is this done? Can you use a similar method to find

$$\begin{pmatrix} 1 & 0 \\ \dfrac{1}{S} & 1 \end{pmatrix}^n ?$$

3. A transmission line, having a series impedance of Z per unit length and a shunt admittance $\left(\text{i.e. } \dfrac{1}{\text{impedance}}\right)$ of Y per unit length may be considered as the limit as $\delta x \to 0$ of a cascade connection of a

number of networks, each corresponding to a length δx of the line, i.e. one unit would be



The corresponding matrix is

$$M = \begin{pmatrix} 1 & Z\dfrac{\delta x}{2} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ Y\delta x & 1 \end{pmatrix} \begin{pmatrix} 1 & Z\dfrac{\delta x}{2} \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 + ZY\dfrac{(\delta x)^2}{2} & Z\delta x + Z^2 Y\dfrac{(\delta x)^3}{4} \\ Y\delta x & 1 + 2Y\dfrac{(\delta x)^2}{2} \end{pmatrix}$$

If l is the length of the line $n\delta x = l$, so that

$$M = \begin{pmatrix} 1 + \dfrac{ZYl^2}{2n^2} & \dfrac{Zl}{n} + \dfrac{Z^2Yl^3}{4n^3} \\ \dfrac{Yl}{n} & 1 + \dfrac{ZYl^2}{2n^2} \end{pmatrix}$$

The problem is then to find M^n, again. There is a 'Binomial Theorem' for matrices (due to Proctor Wilson), i.e.

$$\begin{pmatrix} ch\gamma + ksh\gamma & Z_0(1-k^2)^{\frac{1}{2}}sh\gamma \\ \dfrac{1}{Z_0}(1-k^2)^{\frac{1}{2}}sh\gamma & ch\gamma - ksh\gamma \end{pmatrix}^n$$

$$= \begin{pmatrix} chn\gamma + kshn\gamma & Z_0(1-k^2)^{\frac{1}{2}}shn\gamma \\ \dfrac{1}{Z_0}(1-k^2)^{\frac{1}{2}}shn\gamma & chn\gamma - kshn\gamma \end{pmatrix}$$

Prove it by induction.

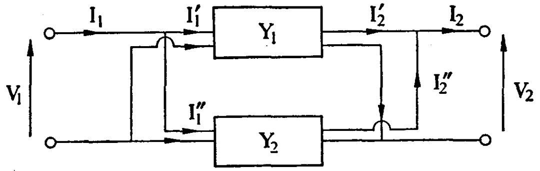
4. The Y matrix is the inverse of the impedance matrix of a circuit Z, where

$$\begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = \begin{pmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{pmatrix} \begin{pmatrix} I_2 \\ I_2 \end{pmatrix} \quad \text{or} \quad V = ZI$$

so

$$\begin{pmatrix} I_1 \\ I_2 \end{pmatrix} = \begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} \quad \text{or} \quad I = YV$$

Consider



Here

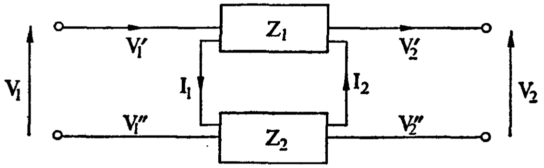
$$\begin{pmatrix} I_1' \\ I_2' \end{pmatrix} = Y_1 \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} I_1'' \\ I_2'' \end{pmatrix} = Y_2 \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$$

But

$$I_1 = I_1' + I_1'' \quad \text{and} \quad I_2 = I_2' + I_2''$$

$$\begin{pmatrix} I_1 \\ I_2 \end{pmatrix} = \begin{pmatrix} I_1' \\ I_2' \end{pmatrix} + \begin{pmatrix} I_1'' \\ I_2'' \end{pmatrix} = (Y_1 + Y_2) \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$$

Prove a similar result using Z matrices for the circuit



For further information see books by J. W. Head and Mary Tropper *Matrix Theory for Electrical Engineering Students* (Harrap), also *Matrix Algebra for Electronic Engineers* by P. Hlawiczka (Iliffe).

VI AFFINE TRANSFORMATIONS

9.34 We have met before the idea of transformations of the plane by the application of matrices. We now consider transformations of the type

$$\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix}$$

or

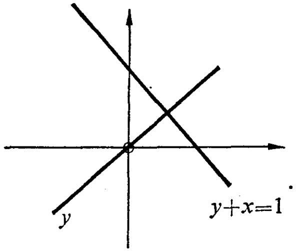
$$\left.\begin{aligned} x &\to ax + by + e \\ y &\to cx + dy + f \end{aligned}\right\}$$

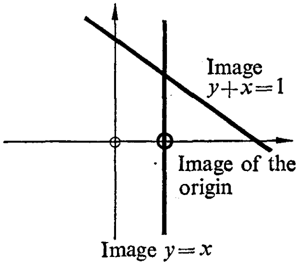
It is easiest, to begin with, to take particular values for a, b, c, d, e and f, and to investigate this particular transformation. Consider

$$\begin{aligned} x &\to x - y + 1 \\ y &\to x + 2y \end{aligned}$$

As with the theory of knots,* our problem is to find out what remains invariant under such a transformation.

We shall find the images of two easy lines which are perpendicular, i.e. $y = x$ and $y + x = 1$.





It is quite clear that straight lines are transformed into straight lines but the angle between them has altered. Are parallel lines mapped into parallel lines? Try $y + x = 1$ and $y + x = 3$. Now experiment for yourself, using the above transformation and answer the following questions.

Exercise 9.8

Investigate whether under the transformation given.

1. Every point has a unique image-point.
2. The image of the intersection of two lines is the intersection of the two image lines.
3. The distance between points is not invariant.
4. The distances between points on the same straight lines are enlarged by the same factor.
5. The midpoint of a line segment is mapped into the midpoint of the image line segment.
6. Circles transform into circles. (Consider points on the circle $x^2 + y^2 = 25$.)
7. Do tangents transform into tangents?
8. Three non-collinear points are mapped into three non-collinear points.

9.35 We have considered a particular transformation, i.e.

$$\begin{pmatrix} x \\ y \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

What problems are associated with a transformation such as

$$\begin{pmatrix} x \\ y \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 1 \\ 6 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix}?$$

---

* See **9.46**.

In fact we have met these problems before when we were investigating the transformations of the plane. To avoid such problems we shall only deal with transformations which are 1–1 and so

$$ad - bc \neq 0$$

is a necessary and sufficient condition for this.

It is interesting to note that it is possible to find a transformation under which any three given points can be mapped into three predetermined points. For example, any triangle can be transformed into an equilateral triangle. For instance, is it possible to find a transformation of the points, (3, 0), (1, 3), (−1, −1) into the points (2, 0), (−2, 0) and $(0, 2\sqrt{3})$? Suppose the transformation is

$$\begin{aligned} x &\rightarrow ax + by + e \\ y &\rightarrow cx + dy + f \end{aligned}$$

Then the image of (3, 0) is $(3a + e, 3c + f)$, and this must be (2, 0). Hence

$$3a + e = 2; \quad 3c + f = 0.$$

Similarly, the image of (1, 3) is $(a + 3b + e, c + 3d + f)$ which must be (−2, 0), so that

$$a + 3b + e = -2; \quad c + 3d + f = 0.$$

Finally the image of (−1, −1) is $(-a - b + e, -c - d + f)$, which is $(0, +2\sqrt{3})$, so

$$-a - b + e = 0; \quad -c - d + f = 2\sqrt{3}$$

We have six equations in six unknowns and so there may be a unique solution. In fact

$$a = \frac{5}{7}; \quad b = -\frac{6}{7}; \quad c = -\frac{3\sqrt{3}}{7}; \quad d = -\frac{2\sqrt{3}}{7}; \quad e = -\frac{1}{7};$$
$$f = \frac{9}{7}\sqrt{3}.$$

In general, then, any three non-collinear points may be transformed into the three vertices of an equilateral triangle.

9.36 It is now clear that it is possible to map a parallelogram into a given square.

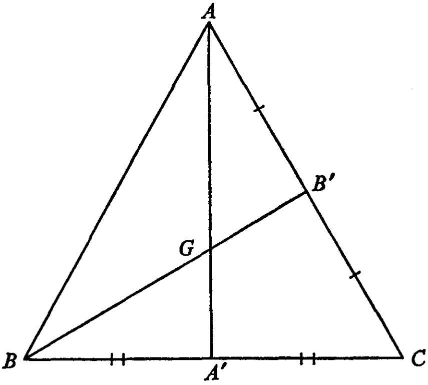
We are now able to prove many properties of triangles and parallelograms by considering them to be mapped into equilateral triangles and squares, proving what we require with these and remembering what properties are invariant under the transformation and then mapping back to our original figures.

This idea is a very important one in electro-statics where we are interested in, for example, the field round an elliptical conductor. The whole problem is transformed into finding the field round a circular conductor, which is much easier, and applying the transformation which will take the circle into the ellipse.

Exercise 9.9

Prove the following, using the ideas of transformations in general.

1. The 'mid-point theorem' for triangles.
2. The medians of a triangle intersect and do so at the point of trisection of the median, i.e. so that here $AG = \frac{2}{3}AA'$; $BG = \frac{2}{3}BB'$.



3. The diagonals of a parallelogram bisect each other.
4. The midpoints of the sides of a parallelogram form the vertices of another parallelogram.
5. Investigate circle theorems and see if you can find corresponding results for the ellipse.

Group structure

9.37 The affine transformations form a group under the rule of 'followed by'.

Closure

Consider the transformation in the matrix form

$$\mathbf{X} \to A\mathbf{X} + \mathbf{B},$$

$$\text{or} \quad \begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix}$$

If we follow this with $\mathbf{X} \to C\mathbf{X} + \mathbf{D}$. Then

$$\begin{aligned}\mathbf{X} \to A\mathbf{X} + \mathbf{B} \to A(C\mathbf{X} + \mathbf{D}) + \mathbf{B} \\ \to (AC)\mathbf{X} + (A\mathbf{D} + \mathbf{B})\end{aligned}$$

which is another affine transformation.

Associativity

This can be demonstrated easily.

Identity

The identity transformation is the one such that,

$$\mathbf{X} \to I\mathbf{X} + \mathbf{O} \quad \text{where} \quad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Inverses

The inverse of

$$\mathbf{X} \to A\mathbf{X} + \mathbf{B}$$

will be

$$A^{-1}(A\mathbf{X} + \mathbf{B} + \mathbf{B}^{-1}) \to X$$

where

$$\mathbf{B}^{-1} = \begin{pmatrix} -e \\ -f \end{pmatrix} \quad \text{and} \quad A^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

9.38 It is clear that the transformations of the type $\mathbf{X} \to A\mathbf{X}$ will be a subgroup of our group of affine transformations. Under what conditions do the affine transformations give us (*a*) isometries (i.e. transformations in which distance is invariant, (*b*) similarities (i.e. transformations in which angles are invariant)?

In fact for isometries

$$\mathbf{X} \to A\mathbf{X} + \mathbf{B}$$

where A is orthogonal, i.e.

$$A \,.\, A' = I$$

(A' is the transpose of A).

For similarities

$$\mathbf{X} \to A\mathbf{X} + \mathbf{B} \quad \text{where} \quad A \,.\, A' = kI$$

Algebraically, for isometries

$$a^2 + b^2 = c^2 + d^2 = 1$$

and

$$ac + bd = 0$$

For similarities

$$a^2 + b^2 = c^2 + d^2 = k \qquad (k > 0)$$

and

$$ac + bd = 0$$

Note that

$$ad - bc \neq 0$$

Check that these are true by taking particular values for a, b, c, d, e and f which satisfy the conditions. For further information on affine transformations see, for example, *A Survey of Modern Algebra* by Birkhoff and MacLane.

VII MATRICES AND THE THEORY OF GAMES

9.39 We shall be considering games of strategy (not of pure chance) between two players, each of whose 'moves' influence but do not completely determine the outcome. Each player must try to make the best possible moves, regardless of what his opponent does, and must try to ensure that his 'payoff' will be as large as possible.

For example, let us call two players R and C (R will have command of the rows and C will have command of the columns in our matrix). R has a silver sixpence, and six copper pennies in a pile (a copper sixpence); C has a silver sixpence and a copper penny.

R has two piles S6 Cu6
C has two piles S6 Cu1

The game is played by exposing the coins simultaneously; if they match in material R wins the positive difference in amount, if they do not match C wins the difference in amount. We can then display the winnings in the form of a matrix.

$$\begin{array}{cc|c|c|}
 & & \multicolumn{2}{c}{\text{Player C}} \\
 & & \multicolumn{1}{c}{\text{S6}} & \multicolumn{1}{c}{\text{Cu1}} \\
\cline{3-4}
\text{Player R} & \text{S6} & 0 & -5 \\
\cline{3-4}
 & \text{Cu6} & 0 & +5 \\
\cline{3-4}
\end{array}$$

The entry -5 means that R pays C fivepence; the entry $+5$ means that C pays R fivepence. The game is completely described by the matrix.

9.40 *Strategies*. Player C would obviously like to achieve the entry -5 by choosing the second column. He then stands to win (or lose) fivepence, depending on R's choice of row. It is obviously sounder to choose column 1 since no matter what R does, C will break even. R on the other hand, would like to achieve $+5$ by choosing the second row.

The optimal strategy then is for player C to play S6 and for player R to play Cu6.

9.41 The *value* of a game is the outcome if both players use the optimal strategy. If the value of the game is zero, then the game is said to be *fair*.

A game is said to be *strictly determined* if the matrix contains an entry v which is simultaneously the *maximum* of its column and the *minimum* of its row. v is the value of the game. Strictly determined games can be solved easily since each player can calculate his opponents best move.

In our example, the game is a fair, strictly determined game. Not all strictly determined games are fair.

Examples

1.

$$\begin{array}{|c|c|}\hline 0 & 3 \\ \hline -2 & 5 \\ \hline \end{array}$$

Fair; strictly determined.

2.

$$\begin{array}{|c|c|}\hline 7 & 4 \\ \hline -3 & -2 \\ \hline \end{array}$$

Strictly determined, value 4

3.

$$\begin{array}{|c|c|}\hline 0 & 3 \\ \hline 2 & 1 \\ \hline \end{array}$$

Not strictly determined.

4.

$$\begin{array}{|c|c|}\hline 2 & 3 \\ \hline 4 & -1 \\ \hline \end{array}$$

Not strictly determined.

Exercise 9.10

Find out which of the following games are strictly determined and give the value of those that are.

1.

$$\begin{array}{|c|c|}\hline 5 & 3 \\ \hline 0 & -2 \\ \hline \end{array}$$

2.

$$\begin{array}{|c|c|}\hline 6 & 0 \\ \hline 0 & 3 \\ \hline \end{array}$$

3.

$$\begin{array}{|c|c|}\hline 2 & -1 \\ \hline -1 & 3 \\ \hline \end{array}$$

4.

$$\begin{array}{|c|c|}\hline 4 & 2 \\ \hline 0 & 5 \\ \hline \end{array}$$

5.

$$\begin{array}{|c|c|} \hline 6 & 0 \\ \hline 0 & -1 \\ \hline \end{array}$$

6.

$$\begin{array}{|c|c|} \hline 2 & 2 \\ \hline 2 & 2 \\ \hline \end{array}$$

9.42 It is quite clear that some 2 × 2 matrix games are not strictly determined, i.e. they have no entry which is simultaneously the *maximum of the column* it is in and also the *minimum of the row* it is in.

Example

Consider a game in which one player hides in his hand either a sixpence or a shilling. If the other player guesses the coin then he wins it, if not he gets nothing. The games matrix is,

$$\begin{array}{cc|c|c|} & & \multicolumn{2}{c}{\text{C guesses}} \\ & & \text{6d} & \text{1s} \\ \cline{3-4} \text{R chooses} & \text{6d} & -\text{6d} & 0 \\ \cline{3-4} & \text{1s} & 0 & -\text{1s} \\ \cline{3-4} \end{array} \quad \text{or in pence} \quad \begin{array}{cc|c|c|} & & \multicolumn{2}{c}{\text{C}} \\ & & 6 & 12 \\ \cline{3-4} \text{R} & 6 & -6 & 0 \\ \cline{3-4} & 12 & 0 & -12 \\ \cline{3-4} \end{array}$$

The game is not strictly determined. How should it be played? If the game is played for a number of times it is clear that R must not play (or hide) the same coin all the time, for C would soon guess this. Nor can R afford to play in any fixed pattern, for once C has realized this then, he can always win.

We now introduce the idea of a *mixed strategy*—for example, for R we say that he is to play row 1 with probability p_1 and row 2 with probability p_2 with $p_1 + p_2 = 1$, $p_1 \geqslant 0$, $p_2 \geqslant 0$. Similarly for player C a mixed strategy will be to play column 1 with probability q_1 and column 2 with probability q_2, with as before $q_1 + q_2 = 1$, $q_1 \geqslant 0$, $q_2 \geqslant 0$.

A mixed strategy vector for R is (p_1, p_2) and for C is $\begin{pmatrix} q_1 \\ q_2 \end{pmatrix}$.

DEFINITION

9.43 For a non-strictly determined game of value v, with *best* strategy vectors of $(p_1\ p_2)$ for R and $\begin{pmatrix} q_1 \\ q_2 \end{pmatrix}$ for C then

(i) $(p_1\ p_2)\begin{pmatrix} a & b \\ c & d \end{pmatrix} \geqslant (v\ \ v)$

(ii) $\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} q_1 \\ q_2 \end{pmatrix} \leqslant \begin{pmatrix} v \\ v \end{pmatrix}$.

where $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is the game matrix. The game is fair if $v = 0$.

If both R and C use their best strategies, equality occurs.

The expectation of R's pay off will be (iii):

$$ap_1q_1 + bp_1q_2 + cp_2q_1 + dp_2q_2 = v$$

(since a occurs with probability p_1q_1 and so on). We can now calculate the values of p_1, p_2, q_1 and q_2 and also v in terms of a, b, c and d, since from (i)

$$p_1a + p_2c = v$$
$$p_1b + p_2d = v$$

from (ii)

$$aq_1 + bq_2 = v$$
$$cq_1 + dq_2 = v$$

and (iii)

$$ap_1q_1 + bp_1q_2 + cp_2q_1 + dp_2q_2 = v$$

Solving (i) remembering that $p_1 + p_2 = 1$,

$$p_1 = \frac{d-c}{a+d-b-c}; \quad p_2 = 1 - p_1$$

and similarly

$$q_1 = \frac{d-b}{a+d-b-c}; \quad q_2 = 1 - q_1$$

and

$$v = \frac{ad-bc}{a+d-b-c}$$

Note that $v = 0$ when $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = 0$.

9.44 First of all, let us consider the previous example (p. 230). The above results give:

$$p_1 = \frac{-12-0}{-6-12-0-0} = \frac{2}{3} \Rightarrow p_2 = \frac{1}{3}$$

$$q_1 = \frac{-12-0}{-6-12-0-0} = \frac{2}{3} \Rightarrow q_2 = \frac{1}{3}$$

$$v = \frac{(-6)\,(-12) - 0.0}{-6-12-0-0}$$

$$= \frac{+72}{-18}$$

$$= -4.$$

This means that R should play row 1 with probability $\frac{2}{3}$, row 2 with probability $\frac{1}{3}$. C should play column 1 with probability $\frac{2}{3}$, column 2 with probability $\frac{1}{3}$. Also the value of the game is given by,

$$-6 \,.\, \tfrac{2}{3} \,.\, \tfrac{2}{3} + 0 \,.\, \tfrac{2}{3} \,.\, \tfrac{1}{3} + 0 \,.\, \tfrac{1}{3} \,.\, \tfrac{2}{3} + -12 \,.\, \tfrac{1}{3} \,.\, \tfrac{1}{3} = -4$$

as before (this is R's expectation).

9.45 Now let us consider the game matrix

$$\begin{array}{|c|c|}\hline 4 & 2 \\ \hline 0 & 5 \\ \hline \end{array}$$

Here

$$\begin{aligned} p_1 &= \frac{5-0}{4+5-2-0} = \frac{5}{7} \Rightarrow p_2 = \frac{2}{7} \\ q_1 &= \frac{5-2}{4+5-2-0} = \frac{3}{7} \Rightarrow q_2 = \frac{4}{7} \\ v &= \frac{4.5-2.0}{4+5-2-0} = \frac{20}{7} \\ &= 2\tfrac{6}{7}. \end{aligned}$$

It is possible to find the probabilities quickly by the following method.

Consider again.

$$\begin{array}{c|c|c|} & \multicolumn{2}{c}{C} \\ \cline{2-3} R & -6 & 0 \\ \cline{2-3} & 0 & -12 \\ \cline{2-3} \end{array}$$

To find the *odds* we proceed as follows. Subtract the numbers in the second row from the corresponding numbers in the first row, to get

$$\begin{array}{|c|c|}\hline -6 & +12 \\ \hline \end{array}$$

We neglect the minus sign and the probabilities for C are as follows. Divide by 12 + 6 to give probabilities

$$\begin{array}{|c|c|}\hline \frac{6}{12+6} & \frac{12}{12+6} \\ \hline \end{array}$$

and then swap them round to give

$$\underset{\text{Probability of playing 1st col.}}{\nearrow} \begin{array}{|c|c|}\hline \frac{2}{3} & \frac{1}{3} \\ \hline \end{array} \underset{\text{Probability of playing 2nd col.}}{\nwarrow}$$

We can do likewise for *R* by subtracting numbers in the second column from the corresponding numbers in the first, changing the odds to probabilities and interchanging as follows:

$$\begin{array}{|c|c|}\hline -6 & 0 \\ \hline 0 & -12 \\ \hline\end{array} \qquad \begin{array}{|c|}\hline -6 \\ \hline +12 \\ \hline\end{array} \qquad \begin{array}{|c|}\hline \frac{6}{12+6} \\ \hline \frac{12}{12+6} \\ \hline\end{array}\Big\updownarrow \qquad \begin{array}{|c|l}\hline \frac{2}{3} & \text{Probability of playing 1st row} \\ \hline \frac{1}{3} & \text{Probability of playing 2nd row} \\ \hline\end{array}$$

It is easy to show that the formulae and this method are equivalent.

Exercise 9.11

Find the best strategies for the following games and the value of each game.

1.

$$\begin{array}{|c|c|}\hline 6 & 4 \\ \hline 1 & 8 \\ \hline\end{array}$$

2.

$$\begin{array}{|c|c|}\hline 5 & 4 \\ \hline 3 & 6 \\ \hline\end{array}$$

3.

$$\begin{array}{|c|c|}\hline 9 & 7 \\ \hline 4 & 11 \\ \hline\end{array}$$

4.

$$\begin{array}{|c|c|}\hline 7 & 3 \\ \hline 11 & 2 \\ \hline\end{array}$$

5.

$$\begin{array}{|c|c|}\hline 1 & 0 \\ \hline -2 & 4 \\ \hline\end{array}$$

6.

$$\begin{array}{|c|c|}\hline 3 & 5 \\ \hline 4 & 6 \\ \hline\end{array}$$

7. *R* has two airfields. He has only enough troops to successfully defend one at a time. *C* can attack one or the other, but not both. One of the airfields is twice as valuable as the other. What strategies should each adopt?

8. A man is rushing home late and he suddenly remembers that it is his wife's birthday—but he is not sure. The only shop open is a tobacconist's where he can buy chocolates.

If he comes home with no chocolates and it is not her birthday, nothing will be said, i.e. pay-off 0. If he comes home with chocolates and it is not her birthday, she will be delighted and offer him a couple of drinks (pay-off 2). If it is her birthday and he has remembered the pay-off will be 3 say! If he has forgotten and it *is* her birthday a miserable evening lies ahead (pay-off −12). What should he do?

For further information on Games Theory and Matrices see *The Compleat Strategyst*, J. D. Williams, McGraw-Hill, 'Teach Yourself Operations Research', and for applications see for example *Quantitative Geography*, Cole and King.

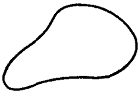
VIII KNOTS

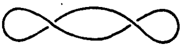
9.46 The elementary theory of knots is an easy example of how matrices can be used to describe various situations. Knots can be divided into two basic types:

(*a*) those which can be regarded as unknotted and are of the same type as the circle.

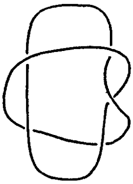
(*b*) those which are knotted.

Knots may be altered by twisting, bending or stretching but not by cutting and rejoining. Under these simple transformations we have to search for an *Invariant* (something which does not change by twisting etc.). This search for an invariant property is a common one in Mathematics and often it is not an easy property to find. What we require, then, is to find a quantity which is unique for each type of knot and remains constant under our simple, allowable transformations, e.g.

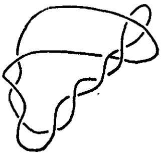




These two diagrams represent an unknotted situation, the second being a loop twisted over. Here is another unknotted situation:



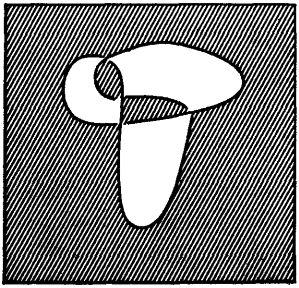
and another



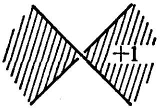
What have such knots in common?

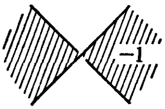
The determinant of a knot

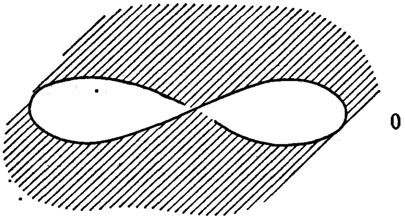
9.47 (*a*) Shade the regions in a chessboard fashion, alternately black and white, with the outside (infinite) area black, e.g.



(*b*) Give a value to each cross-over point, either +1, −1 or 0 as shown.







(This value is given when the two shaded areas are, in fact, the same.)

(*c*) We now number the shaded regions (except for the infinite one).

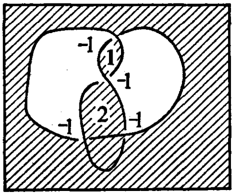
(*d*) Form the matrix K with entries as follows:

k_{pp} = the sum of the values on the area labelled p,

$k_{pq} = k_{qp}$ = minus the sum of the values common to regions p and q.

(*e*) The numerical value of the determinant of K is a *knot invariant*.

9.48 *For example*, consider the knot below:

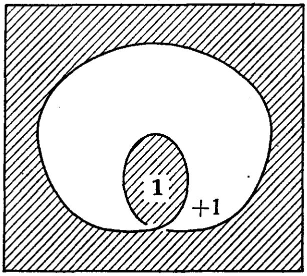


The matrix

$$K = \begin{pmatrix} -2 & +1 \\ +1 & -3 \end{pmatrix}$$

and so det $K = +5$.

NOTE. For a circle, clearly of the same form as the invariant is +1.

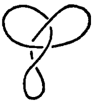


This is a rough and ready method for determining the type of the knot. There is a much more powerful method as we shall see.

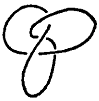
Exercise 9.12

Find the knot invariant in the following cases:

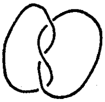
1.



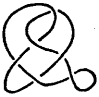
2.



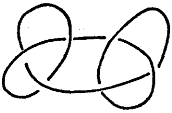
3.



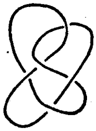
4.



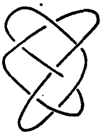
5.



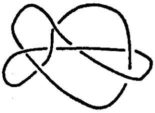
6.



7.



8.

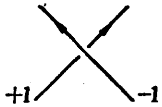


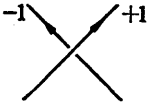
A more powerful method

9.49 (*a*) Orient the knot by following round the knot and putting in direction arrows.

(*b*) Label each cross-over point and each piece of the knot (i.e. between going under one part to going under another).

(*c*) Give the value +1 and −1 to the pieces at a cross-over point.





(*d*) Find the matrix K with rows corresponding to cross-over points and columns corresponding to pieces.

(*e*) Construct L by deleting last row and column of K.

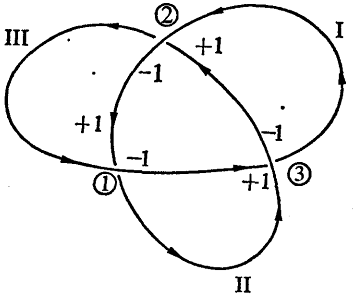
(*f*) Construct M from L by adding first column to second, second to third and so on.

(*g*) For any value of $n \in Z^+$, the set of positive integers, the numerical value of the determinant of the matrix

$$M^n - (M - I)^n$$

is another knot invariant.

9.50 *For example,* consider the knot



Then

$$K = \begin{array}{cc} & \begin{array}{ccc} \text{I} & \text{II} & \text{III} \end{array} \\ \begin{array}{c} \textcircled{1} \\ \textcircled{2} \\ \textcircled{3} \end{array} & \begin{pmatrix} +1 & 0 & -1 \\ -1 & +1 & 0 \\ 0 & -1 & +1 \end{pmatrix} \end{array}$$

$$L = \begin{pmatrix} +1 & 0 \\ -1 & +1 \end{pmatrix}, \qquad M = \begin{pmatrix} +1 & +1 \\ -1 & 0 \end{pmatrix}.$$

Show that for

$n = 1$ the knot invariant is 1 (always)
$n = 2$ the knot invariant is 3 (same as by first method)
$n = 3$ the knot invariant is 4
$n = 4$ the knot invariant is 3
$n = 5$ the knot invariant is 1
$n = 6$ the knot invariant is 0
$n = 7$ the knot invariant is 1 and the pattern repeats.

The values given for $n = 2$ and $n = 3$ enable us to distinguish the different types of knot apart from reef knots and granny knots.

Exercise 9.13

Choose one of the first three questions in the previous exercise and calculate this new knot invariant.

For further work on knots see articles in the Scientific American, and books, e.g. *Introduction to Knot theory*—Cromwell and Fox, Ginn 1963.

IX MATRICES WITH COMPLEX ELEMENTS*

9.51 A matrix with real elements can have three main operations performed on it: negativizing, transposing and inverting. The first of these consists of just multiplying each element of the matrix by -1: it is written $-A$. The second consists of writing the rows as the columns of the transpose matrix, and the columns as the rows: thus an $m \times n$ matrix becomes $n \times m$ when transposed. It is written A'. The third operation consists of finding a matrix A^{-1} such that $AA^{-1} = I$, where I is the unit matrix for the set of matrices concerned.

If a matrix has complex elements, we may add a fourth operation to these three. This operation consists of replacing each element of the matrix by its complex conjugate, to give a new matrix $\bar{A}$.

9.52 Consider an Argand plane with complex axes u and v. These can be transformed by a matrix

$$Q = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

where α, β, γ and δ may be complex. Then if the transformed axes are u', v' we have

$$\begin{aligned} u' &= \alpha u + \beta v \\ v' &= \gamma u + \delta v \end{aligned}$$

We restrict this matrix Q to be a unitary matrix of determinant 1. A unitary matrix is a matrix with the special property that $A(\bar{A}') = I$, or $(\bar{A}')^{-1} = A$. If $\bar{\alpha}$, $\bar{\beta}$, $\bar{\gamma}$ and $\bar{\delta}$ are the complex conjugates of the elements of Q, then

$$Q(\bar{Q}') = I$$

where $\bar{Q}$ is the matrix which has the complex conjugate entries corresponding to the entries in Q, i.e. if

$$Q = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \quad \bar{Q} = \begin{pmatrix} \bar{\alpha} & \bar{\beta} \\ \bar{\gamma} & \bar{\delta} \end{pmatrix}$$

---

* We are indebted to Andrew Ellis for this section.

Now
$$\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}\begin{pmatrix}\bar{\alpha} & \bar{\gamma}\\ \bar{\beta} & \bar{\delta}\end{pmatrix} = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$$

$$\alpha\bar{\alpha} + \beta\bar{\beta} = 1 \tag{1}$$

$$\alpha\bar{\gamma} + \beta\bar{\delta} = 0 \tag{2}$$

$$\bar{\alpha}\gamma + \bar{\beta}\delta = 0 \tag{3}$$

$$\gamma\bar{\gamma} + \delta\bar{\delta} = 1 \tag{4}$$

The second and third equations are equivalent, so

$$\alpha\bar{\alpha} + \beta\bar{\beta} = \gamma\bar{\gamma} + \delta\bar{\delta} = 1 \tag{5}$$

$$\bar{\alpha}\alpha + \bar{\beta}\beta = 0 \tag{6}$$

Equation (5) is an equation in real numbers, since $\alpha\bar{\alpha}$, $\beta\bar{\beta}$, $\gamma\bar{\gamma}$, $\delta\bar{\delta}$ are all real, Equation (6), $\bar{\alpha}\gamma + \bar{\beta}\delta = 0$, contains complex numbers, and so we have two conditions imposed. We have also $\alpha\delta - \beta\gamma = 1$ (since $\Delta = 1$). All told there are five conditions. We usually leave the matrix Q as $\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}$, although we may express it in terms of 3 independent variables.

9.53 Consider for example a matrix

$$P = \begin{pmatrix} z & x - jy\\ x + jy & -z\end{pmatrix},$$

where x, y, z may be taken as the coordinates of a point in 3 dimensions, or as any 3 real ordered numbers. Note that

$$\bar{P}' = \begin{pmatrix} z & x + jy\\ x - jy & -z\end{pmatrix}' = \begin{pmatrix} z & x - jy\\ x + jy & -z\end{pmatrix} = P \tag{7}$$

A matrix with this property, $P = (\bar{P}')$, is called *self-adjoint* or *hermitian*. The matrix P also has the property that its trace (i.e. the sum of the elements along the main diagonal) is zero; in this case

$$-z + z = 0 \tag{8}$$

Let us transform the matrix P as follows, to give

$$P_1 = QP\bar{Q}' = QPQ^{-1} \quad \text{since} \quad \bar{Q}' = Q^{-1}$$

The two properties (7) and (8) mentioned above are invariant under transformations of this type. Hence the new matrix P_1 may be written

$$P_1 = \begin{pmatrix} z_1 & x_1 - jy_1\\ x_1 + jy_1 & -z_1\end{pmatrix}$$

Also, as the determinant of Q is 1, the determinant of Q^{-1} is also 1, and so

$$\begin{aligned}\Delta P_1 &= \Delta Q \,.\, \Delta P \,.\, \Delta Q^{-1}\\ \Delta P_1 &= \Delta P\end{aligned}$$

i.e.

$$\begin{aligned}-z^2 - x^2 - y^2 &= -z_1^2 - x_1^2 - y_1^2\\ x^2 + y^2 + z^2 &= x_1^2 + y_1^2 + z_1^2 = \Delta P = \Delta P_1.\end{aligned}$$

If x, y, z are coordinates, then the length of a general vector

$$\mathbf{h} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}}$$

is unaltered by any transformation with this property. The transformation is a rotation of some sort, and can be represented by a 3 × 3 real orthogonal matrix, since an orthogonal matrix is defined as one for which the transpose equals the inverse, or

$$A = (A')^{-1}$$

The unitary 2 × 2 complex matrix Q corresponds to a real 3 × 3 orthogonal matrix. This matrix is found as follows:

9.54 Let D be a 3 × 3 orthogonal matrix, transforming (x, y, z) to (x_1, y_1, z_1). Then $\mathbf{h}' = D\mathbf{h}$. Let the corresponding 2 × 2 unitary matrix be Q_1: then

$$P_1 = Q_1P(\bar{Q}_1').$$

With another 3 × 3 orthogonal matrix B, transforming (x_1, y_1, z_1) to (x_2, y_2, z_2), the corresponding Q-matrix is Q_2. Then

$$\begin{aligned}\mathbf{h}'' &= B\mathbf{h}'\\ &= BD\mathbf{h}\end{aligned}$$

And

$$P_2 = Q_2P_1(\bar{Q}_2')$$

Now let $C = BD$. There is a matrix Q_3 corresponding to C. But

$$P_2 = Q_2P_1(\bar{Q}_2') = Q_2Q_1P(\bar{Q}_1')(\bar{Q}_2')$$

Since $(\bar{Q}_1'\bar{Q}_2') = (\overline{Q_2Q_1})'$, we have $Q_3 = Q_2Q_1$ as the matrix corresponding to C. In fact, all the relations satisfied by the matrices of one set (such as D) are satisfied equally well by the corresponding matrices of the other set (such as Q): the two sets of matrices can be shown to be isomorphic. Now

$$(\bar{Q}') = Q^{-1} = \begin{pmatrix}\delta & -\beta\\ -\gamma & \alpha\end{pmatrix} \quad \text{as} \quad \Delta Q = 1$$

and thus

$$P_1 = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} z & x - jy \\ x + jy & -z \end{pmatrix} \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix}$$

$$= \begin{pmatrix} \alpha z + \beta(x + jy) & \alpha(x - jy) - \beta z \\ \gamma z + \delta(x + jy) & \delta(x - jy) - \delta z \end{pmatrix} \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix}$$

$$= \begin{pmatrix} (\alpha\delta + \beta\gamma)z + \beta\delta(x + jy) - \alpha\gamma(x - jy) & -2\alpha\beta z + \alpha^2(x - jy) - \beta^2(x + jy) \\ 2\gamma\delta z - \gamma^2(x - jy) + \delta^2(x + jy) & -(\alpha\delta + \beta\gamma)z - \beta\delta(x + jy) + \alpha\gamma(x - jy) \end{pmatrix}$$

Thus

$$\begin{aligned} z_1 &= (\alpha\delta + \beta\gamma)z + \beta\delta(x + jy) - \alpha\gamma(x - jy) \\ x_1 + jy_1 &= 2\gamma\delta z - \gamma^2(x - jy) + \delta^2(x + jy) \\ x_1 - jy_1 &= -2\alpha\beta z + \alpha^2(x - jy) - \beta^2(x + jy) \end{aligned}$$

For a general 3 × 3 matrix, with entries a_{ij}

$$z_1 = a_{31}x + a_{32}y + a_{33}z$$

giving

$$z_1 = (\beta\delta - \alpha\gamma)x + j(\alpha\gamma + \beta\delta)y + (\alpha\delta + \beta\gamma)z$$

and the 3 × 3 real orthogonal matrix is

$$\begin{pmatrix} \frac{1}{2}(\alpha^2 - \gamma^2 + \delta^2 - \beta^2) & \frac{1}{2}j(\gamma^2 - \alpha^2 + \delta^2 - \beta^2) & \gamma\delta - \alpha\beta \\ \frac{1}{2}j(\alpha^2 + \gamma^2 - \beta^2 - \delta^2) & \frac{1}{2}(\alpha^2 + \gamma^2 + \beta^2 + \delta^2) & -j(\alpha\beta + \delta\gamma) \\ \beta\delta - \alpha\gamma & j(\alpha\gamma + \beta\delta) & \alpha\delta + \beta\gamma \end{pmatrix}$$

This is the *D*-matrix corresponding to our *Q*-matrix.

9.55 The position of a body in 3 dimensions may be specified by (x, y, z). By means of the matrix above a transformation may be applied. It is equally true that the position of a body in space may be specified by the four parameters α, β, γ and δ, which are called the Cayley-Klein parameters.

The original matrix P which we used in these transformations is

$$x\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + y\begin{pmatrix} 0 & -j \\ j & 0 \end{pmatrix} + z\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

These three matrices above are called σ_x, σ_y and σ_z. They are the ***Pauli spin matrices***, and are associated with the spin of an electron in orbit. These matrices, together with the unit matrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, form a set of four independent matrices, from which any 2 × 2 matrix can be built.

In the complex space considered, the matrix corresponding to a rotation through an angle ϕ about the z-axis is

$$\begin{pmatrix} e^{j\phi/2} & 0 \\ 0 & e^{-j\phi/2} \end{pmatrix}$$

9.56 If we consider a rotation through the angle 2π about the z-axis, in space of 3 dimensions we shall return to the original set of axes. But if in the above matrix

$$\phi = 2\pi$$

then

$$\begin{pmatrix} e^{j\pi} & 0 \\ 0 & e^{-j\pi} \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

This matrix *must* correspond to the 3×3 unit matrix. In fact, there are two 2×2 complex unitary matrices isomorphic to any given 3×3 real orthogonal matrix (one being minus the other). Remember that the 2-dimensional complex space constructed here is an entirely mathematical concept, and so does not require the same properties as 3-dimensional space.

9.57 A vector in the complex space is called a *spinor*. Although the space is a purely mathematical conception, it is used in the realm of quantum mechanics. The wave function of an electromagnetic radiation must be made a spinor in order to include the effects of electron spin.

References

H. GOLDSTEIN, *Classical Mechanics.*

H. W. TURNBULL and A. C. AITKEN, *Introduction to the Theory of Canonical Matrices.*

Answers

Exercise 1.1

1. (*a*) 2×2 (*b*) 3×1 (*c*) 1×1 (*d*) 1×2 *e*) 3×2
(*f*) 2×3 (*g*) 1×4 (*h*) 2×1 (*i*) 4×3

2. Follow since the addition of real numbers is commutative and associative.

4. *0.*

Exercise 1.2

1. (*a*) (32) (*b*) $\begin{pmatrix} 4 & 8 & 12 \\ 5 & 10 & 15 \\ 6 & 12 & 18 \end{pmatrix}$ (*c*) Does not exist.
(*d*) Does not exist (*e*) Does not exist (*f*) $\begin{pmatrix} 4x + 5y + 6z \\ 4p + 5q + 6r \end{pmatrix}$

2. In general no.

3. $n = p$, $m = q$.

5. True.

6. A, B must be square and of the same order: further conditions must also be imposed on the elements.

7. For example $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

Exercise 1.3

3. $R = \left(\frac{b}{a + b}, \frac{a}{a + b}\right)$. In this case $\left(\frac{10}{19}, \frac{9}{19}\right)$

Exercise 1.4

1. (*a*) -2 (*b*) 0 (*c*) $+236$
(*d*) $a^2 - b^2$ (*e*) -8 (*f*) -23
(*g*) 0 (*h*) *abc*

5. (*a*) 43 (*b*) -43

6. (*c*) (i) -430 (ii) -258 (iii) -43

7. (*c*) (i) 0 (ii) -7

Exercise 1.5

1. (i) -74 (ii) $-10{,}636$ (iii) 7000
(iv) 448 (v) 12 (vi) 0

3. 1 and approximately 16·5, 1·5.

4. 0.

6. (*b*) (i) 24 (ii) 60 (iii) *adf* (iv) $\sqrt{30}$

7. 2^n.

8. $|A| = 98$; $|A^2| = (98)^2$; $|5A| = 125|A|$; $5|A| = 5.98$; $|\frac{1}{3}A'| = \frac{1}{27}|A|$;
$|A^3A'| = (98)^4$.

9. (*a*) $(p - q)(q - r)(r - p)$ (*b*) $(p - q)(q - r)(r - p)(p + q + r)$

11. (ii) $-(x - y)(y - z)(z - x)(x + y + z)(x^2 + y^2 + z^2)$

12. -22 13. $3\omega^2$ 14. A, B, C are collinear.

16. (i) $x = a, 1, \frac{-a}{1 + a}$ (ii) $(x - y)^{n-1}\{x + (n - 1)y\}$

17. (i) $(a+b+c+d)(a-b+c-d)((a-c)^2+(b-d)^2)$
(ii) $(x+y+z)^2(x^2+y^2+z^2-xy-yx-zx)^2$

18. (i) $-4(b-c)(c-a)(a-b)$
(ii) $(x+a+b+c)(x+a-b-c)(c-b+a-x)(b-c+a-x)$

Exercise 1.6

2. (*a*) $\begin{pmatrix} \frac{5}{13} & -\frac{4}{13} \\ -\frac{3}{13} & \frac{5}{13} \end{pmatrix}$ (*b*) $\begin{pmatrix} -\frac{1}{4} & -\frac{5}{4} & \frac{9}{4} \\ \frac{1}{4} & \frac{3}{4} & -\frac{5}{4} \\ -\frac{1}{4} & \frac{1}{4} & -\frac{1}{4} \end{pmatrix}$

(*c*) $\frac{1}{9}\begin{pmatrix} 6 & 19 & -13 \\ -3 & -3 & 3 \\ 0 & -7 & 4 \end{pmatrix}$ (*d*) $\begin{pmatrix} \frac{1}{10} & \frac{9}{10} & -\frac{17}{15} \\ -\frac{1}{10} & \frac{11}{10} & -\frac{6}{5} \\ \frac{1}{10} & -\frac{21}{10} & \frac{38}{15} \end{pmatrix}$

(*e*) $\begin{pmatrix} 2 & -3 & 1 \\ 3 & -5 & 1 \\ -5 & 9 & -2 \end{pmatrix}$ (*f*) $-\dfrac{1}{213}\begin{pmatrix} 1 & -38 & 16 \\ -51 & 21 & 36 \\ 11 & 8 & -37 \end{pmatrix}$

3. $AC_2 - I = \begin{pmatrix} -\frac{1}{8} & \frac{3}{4} & 0 \\ 0 & -\frac{1}{8} & 0 \\ 0 & -\frac{7}{16} & -\frac{1}{64} \end{pmatrix}$ $C_1 = \begin{pmatrix} \frac{3}{2} & -1 & 0 \\ 0 & \frac{3}{2} & 0 \\ 0 & 1 & \frac{5}{4} \end{pmatrix}$

$C_2 = \begin{pmatrix} \frac{7}{4} & -2 & 0 \\ 0 & \frac{7}{4} & 0 \\ 0 & \frac{7}{4} & \frac{21}{16} \end{pmatrix}$

4. $B' = \begin{pmatrix} 2 & -2 & 1 \\ 2 & 1 & -2 \\ 1 & 2 & 2 \end{pmatrix}$ $BB' = \begin{pmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{pmatrix}$

$B^{-1} = \begin{pmatrix} \frac{2}{9} & -\frac{2}{9} & \frac{1}{9} \\ \frac{2}{9} & \frac{1}{9} & -\frac{2}{9} \\ \frac{1}{9} & \frac{2}{9} & \frac{2}{9} \end{pmatrix}$ $C = \begin{pmatrix} \frac{1}{9} & -\frac{4}{9} & \frac{8}{9} \\ \frac{4}{9} & -\frac{7}{9} & -\frac{4}{9} \\ \frac{8}{9} & \frac{4}{9} & \frac{1}{9} \end{pmatrix}$

Exercise 2.2

1. $\begin{pmatrix} 3 & -4 \\ -2 & 3 \end{pmatrix}$ 2. $\begin{pmatrix} 2 & -3 \\ -1 & 1 \end{pmatrix}$ 3. $\begin{pmatrix} \frac{11}{2} & -\frac{13}{2} \\ -5 & 6 \end{pmatrix}$

4. $\begin{pmatrix} \dfrac{a}{a^2-b^2} & \dfrac{-b}{a^2-b^2} \\ \dfrac{-b}{a^2-b^2} & \dfrac{a}{a^2-b^2} \end{pmatrix}$ 5. $X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

Exercise 2.3

1. $\begin{pmatrix} 1 & -2 \\ -1 & 3 \end{pmatrix}$ 2. $\begin{pmatrix} 2 & -3 \\ -5 & 8 \end{pmatrix}$ 3. $\begin{pmatrix} -11 & 10 \\ 10 & -9 \end{pmatrix}$

4. $\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}$ 5. $\begin{pmatrix} -\frac{5}{7} & \frac{6}{7} \\ \frac{3}{7} & -\frac{5}{7} \end{pmatrix}$ 6. $\dfrac{1}{ps-qr}\begin{pmatrix} s & -q \\ -r & p \end{pmatrix}$

Exercise 2.4

2. (i) $\begin{pmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 7 & 8 & 9 \end{pmatrix}$ (ii) $\begin{pmatrix} 3 & 6 & 9 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ (iii) $\begin{pmatrix} \frac{10}{3} & \frac{14}{3} & 6 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$

(iv) $\begin{pmatrix} 1 & 2 & 3 \\ 14 & 25 & 36 \\ 7 & 8 & 9 \end{pmatrix}$ (v) $\begin{pmatrix} 1 & 2 & 3 \\ 2 & \frac{5}{2} & 3 \\ 7 & 8 & 9 \end{pmatrix}$ (vi) $\begin{pmatrix} 1 & 2 & 3 \\ 18 & 21 & 24 \\ 7 & 8 & 9 \end{pmatrix}$

Exercise 2.5

1. $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$ 2. $\begin{pmatrix} \dfrac{1}{k} & 0 & 0 \\ 0 & \dfrac{1}{l} & 0 \\ 0 & 0 & \dfrac{1}{m} \end{pmatrix}$

3. $\begin{pmatrix} 1 & -k & km-l \\ 0 & 1 & -m \\ 0 & 0 & 1 \end{pmatrix}$

4. $\begin{pmatrix} 1 & 0 & 0 \\ 4 & -1 & 5 \\ 3 & 0 & -1 \end{pmatrix}$

5. $\begin{pmatrix} \dfrac{1}{a} & 0 & 0 \\ -\dfrac{f}{ab} & \dfrac{1}{b} & 0 \\ \dfrac{-e}{ac}+\dfrac{df}{abc} & \dfrac{-d}{bc} & \dfrac{1}{c} \end{pmatrix}$

6. $\begin{pmatrix} 1 & 4 & 2 \\ 2 & 9 & 4 \\ 2 & 8 & 5 \end{pmatrix}$

7. $\begin{pmatrix} 29 & 2 & 18 \\ 14 & 1 & 9 \\ 14 & 1 & 10 \end{pmatrix}$

8. Singular

Exercise 2.9

1. $\begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

2. $\begin{pmatrix} 1 & 2 & 2 \\ 4 & 9 & 8 \\ 2 & 4 & 5 \end{pmatrix}$

3. $\begin{pmatrix} \frac{1}{18} & \frac{1}{9} & \frac{1}{18} \\ \frac{31}{18} & \frac{7}{9} & \frac{1}{18} \\ \frac{13}{9} & \frac{5}{9} & \frac{1}{9} \end{pmatrix}$

4. $\begin{pmatrix} -\frac{13}{3} & \frac{19}{8} & 2 \\ 1 & -1 & -1 \\ \frac{4}{3} & -\frac{7}{3} & 0 \end{pmatrix}$

5. $\begin{pmatrix} -3 & 9 & -5 \\ 2 & -5 & 3 \\ 1 & -2 & 1 \end{pmatrix}$

6. $\begin{pmatrix} -\frac{5}{2} & 1 & \frac{1}{2} \\ 1 & 1 & -1 \\ \frac{1}{2} & -1 & \frac{1}{2} \end{pmatrix}$

7. $\begin{pmatrix} 1 & -1 & 0 & 0 \\ -3 & 4 & 3 & -2 \\ 2 & -2 & -2 & 1 \\ 0 & 1 & -2 & 0 \end{pmatrix}$

8. $\begin{pmatrix} 7 & -2 & -2 & -2 \\ -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}$

9. $\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 & -1 \\ 2 & 0 & 0 & -2 \end{pmatrix}$

10. $\begin{pmatrix} 7 & -3 & -3 & -3 \\ -3 & 2 & 1 & 1 \\ -3 & 1 & 2 & 1 \\ -3 & 1 & 1 & 2 \end{pmatrix}$

Exercise 3.1

1. $x = 1, y = 2$
2. $x = 1, y = -1, z = 4$
3. $x = -2,\ y = 3, z = 4$
4. $\Delta = 0$

Exercise 3.2

1. (i) $x = -\frac{7}{2}, y = 3, z = \frac{7}{2}$ (ii) $x = -2, y = 4, z = 0$

2. (i) $x = t, y = \dfrac{36-5t}{3}, z = \dfrac{2t+12}{3}$ (iii) Three identical equations.

Exercise 3.3

1. (*a*) $x = 2, y = 1$ (*b*) $x = \frac{5}{16}, y = -\frac{45}{16}$
   (*c*) inconsistent (*d*) same equation twice
   (*e*) $x = -3, y = -3$ (*f*) $x = -27, y = 25$

2. Yes. 3. (2, 1) 4. $x = 1, y = 1$

6. $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}; x = 1, y = 1$

7. $\begin{vmatrix} a & b & p \\ c & d & q \\ e & f & r \end{vmatrix} = 0$

Exercise 3.4

1. $x = 2, y = 3$
2. No solution
3. $x = -2, y = -4$
4. $x = t, y = \frac{4}{3}t - \frac{2}{3}$
5. $x = \frac{13}{22}, y = -\frac{1}{22}$
6. No solution
7. No solution

Exercise 3.5

1. $x = \frac{9}{14}, y = \frac{2}{7}, z = \frac{31}{14}$

2. $x = 1, y = 2, z = 3$

3. $x = t, y = \dfrac{6 - 3t}{2}, z = \dfrac{24 - 13t}{2}$

4. $x = \dfrac{24 - 2t}{5}, y = \dfrac{3 - 3t}{13}, z = t$

5. $x = \frac{9}{13}, y = \frac{1}{7}, z = \frac{15}{13}$

6. $x = 1, y = 0, z = -57$

7. $x = 0, y = 0, z = 1$

8. $x = 1, y = 10, z = 1$

9. $x = \dfrac{225}{117}, y = \dfrac{24}{39}, z = 0$

10. $x = t, y = \dfrac{47 - 64t}{14}, z = \dfrac{11}{2} - 10t$

11. $x = -2, y = 3\frac{1}{4}, z = 8\frac{1}{4}$

Exercise 3.6

1. (*a*) No solution

(*b*) Three parallel planes

(*c*) $x = 1, y = -1, z = 2$

(*d*) $x = t, y = \dfrac{34 - 31t}{9}, z = \dfrac{10 - 7t}{9}$

(*e*) $x = t, y = \dfrac{-6 - t}{6}, z = \dfrac{6 - 5t}{6}$

(*f*) No solution

(*g*) $x = 3, y = 2, z = 1$

(*h*) $x = t, y = \dfrac{20 + 6t}{33}, z = \dfrac{-8 - 9t}{11}$

2. $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$; $x = 113{\cdot}7, y = -76{\cdot}8, z = 81{\cdot}15$

4. (ii) $x = 1, y = 2, z = 3$.

5. $\begin{pmatrix} 1 & 0 & 1 \\ -14 & 4 & -17 \\ 11 & -3 & 13 \end{pmatrix}$; $x = 1, y = -8, z = 7$; $x = 4, y = -69, z = 53$

6. $x = 1, y = 2, z = -1$

7. $x = 10, y = 9, z = 11$ (ii) and (iii) No solutions

8. $a = 3$

9. $\begin{pmatrix} -\frac{6}{5} & \frac{1}{10} & \frac{17}{10} \\ \frac{1}{5} & -\frac{1}{10} & \frac{8}{10} \\ \frac{4}{5} & \frac{1}{10} & -\frac{13}{10} \end{pmatrix}$; $x = 2, y = 1, z = -2$

10. $x_1 = \frac{2}{7}, x_2 = -\frac{1}{7}, x_3 = \frac{1}{7}$

11. $x = 1, y = -2, z = 3$

Exercise 3.7

1. $\lambda = 1$, or $\lambda = 5$

2. $x = t, y = -4t, z = 2t$

3. $-0{\cdot}84$

Exercise 3.8

1. $x = 1, y = 1, z = 0$

2. $x = 2, y = -1, z = 3$.

Exercise 3.9

1. $x = 2, y = 1, z = -2$

2. $x = \dfrac{18 + 7t}{11}, y = \dfrac{8 + 19t}{11}, z = t$

3. Same equation 3 times

4. No solution

Exercise 3.10

1. $(0, 0, 0), (\frac{1}{4}, \frac{1}{8}, \frac{5}{8}), (\frac{11}{6}, -\frac{2}{3}, -\frac{1}{6}), (5, -7, 3)$

2. $x = -(\alpha + \beta + \gamma), y = 1, z = 1$. 3 identical planes

3. $x = \dfrac{1}{\Delta}(3ab - 8a^2 - 5b + 5ac + 8 - 3c)$

$y = \dfrac{1}{\Delta}(2b - 2ac - ab + 3a^2 + c - 3), z = \dfrac{1}{\Delta}(c - a - 1)$

where $\Delta = b - a - 1$.

4. $a \neq b \neq c$, $x = \dfrac{-bc}{(a-b)(c-a)}$, $y = \dfrac{ac}{(a-b)(b-c)}$, $z = \dfrac{ab}{(a-c)(b-c)}$

5. (i) $x = \frac{1}{3}$, $y = -\frac{1}{2}$, $z = \frac{5}{6}$ (ii) $x = t$, $y = \dfrac{-5-3t}{12}$, $z = \dfrac{33t-1}{12}$

6. $x = \dfrac{b-1}{a(ab-1)}$, $y = \dfrac{b(a^3b - a^2 - ab + 1)}{(ab-1)^2}$, $z = \dfrac{b(1-a)(a^2+b-a-1)}{(ab-1)^2}$

7. $k = 1$

8. $x = \dfrac{1}{\Delta} d(2b+c)$, $y = \dfrac{-d}{\Delta}(2a+c)$, $z = \dfrac{d}{\Delta}(b-a)$

   where Δ is the determinant of the coefficients.

9. $3abc - a^3 - b^3 - c^3 = 0$; $a + b + c = 0$; $\alpha + \beta + \gamma = 0$; $a = b = c$; $\alpha = \beta = \gamma$

10. $X = 3kg$, $Y = -\frac{1}{4}kg$, $Z = 5\frac{1}{4}kg$

Exercise 3.11

1. (*a*) $x + y = 1$ (*b*) $y = 1$ (*c*) $\dfrac{x}{a} + \dfrac{y}{b} = 1$

2. (*a*) $x^2 + y^2 - 6x + 8y = 20$ (*b*) $x^2 + y^2 - 6x - 8y = 0$
   (*c*) Not possible; the given points are collinear.
   (*d*) $x^2 + y^2 - 6x - 8y = 0$

3. (*a*) $6x + 4y + 3z = 12$
   (*b*) Not possible by this method. The given points are collinear, and satisfy the relation $x - 2 = \dfrac{y+3}{2} = \dfrac{z-4}{2}$. There are an infinite number of possible planes, but none are determined by this method.
   (*c*) $3x - 4y + z = 36$

4. $(y-h)^2 = 4a(x-k)$

$$\begin{vmatrix} y^2 & x & y & 1 \\ y_1^2 & x_1 & y_1 & 1 \\ y_2^2 & x_2 & y_2 & 1 \\ y_3^2 & x_3 & y_3 & 1 \end{vmatrix} = 0; \ (y+1)^2 = -16x$$

5.
$$\begin{vmatrix} x^2 & x & y & 1 \\ x_1^2 & x_1 & y_1 & 1 \\ x_2^2 & x_2 & y_2 & 1 \\ x_3^2 & x_3 & y_3 & 1 \end{vmatrix} = 0; \ (y-4) = (x+1)^2$$

6.
$$\begin{vmatrix} x^2 + y^2 + z^2 & x & y & z & 1 \\ x_1^2 + y_1^2 + z_1^2 & x_1 & y_1 & z_1 & 1 \\ x_2^2 + y_2^2 + z_2^2 & x_2 & y_2 & z_2 & 1 \\ x_3^2 + y_3^2 + z_3^2 & x_3 & y_3 & z_3 & 1 \\ x_4^2 + y_4^2 + z_4^2 & x_4 & y_4 & z_4 & 1 \end{vmatrix} = 0$$

If $\begin{vmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \\ x_4 & y_4 & z_4 & 1 \end{vmatrix} = 0$, then the four points are coplanar.

Note that this does not necessarily mean there is no sphere through the points; if the four points lie on a circle in this plane then an infinite number of spheres could be found, but none are determined by the above equation.

7. $\dfrac{x^2}{9} + \dfrac{y^2}{16} - \dfrac{z^2}{144} = 1$

8. (*a*) $\dfrac{x^2}{24^2} - \dfrac{y^2}{12^2} = 1$, a hyperbola (*b*) $\dfrac{x^2}{25^2} + \dfrac{y^2}{100^2} = 1$, an ellipse
   (*c*) $x^2 - y^2 = 9$, a rectangular hyperbola
   (*d*) $(x + 2y)(x - 2y) = 0$, a line pair
   (*e*) $x^2 + y^2 = 25^2$, a circle

9. (*a*) 5 points determine a conic

$$\begin{vmatrix} x^2 & xy & y^2 & x & y & 1 \\ x_1^2 & x_1y_1 & y_1^2 & x_1 & y_1 & 1 \\ x_2^2 & x_2y_2 & y_2^2 & x_2 & y_2 & 1 \\ x_3^2 & x_3y_3 & y_3^2 & x_3 & y_3 & 1 \\ x_4^2 & x_4y_4 & y_4^2 & x_4 & y_4 & 1 \\ x_5^2 & x_5y_5 & y_5^2 & x_5 & y_5 & 1 \end{vmatrix} = 0$$

(*b*) $xy + y^2 - 3x = 9$, a hyperbola.

10. Nine points determine a cubic curve.
The top row of the required determinant is

$$x^3 \quad x^2y \quad xy^2 \quad y^3 \quad x^2 \quad xy \quad y^2 \quad x \quad y \quad 1$$

11. Nine points determine a quadric surface.
The top of the required determinant is

$$x^2 \quad y^2 \quad z^2 \quad xy \quad yz \quad zx \quad x \quad y \quad z \quad 1$$

12.

$$\begin{vmatrix} x & y & z & t & 1 \\ x_1 & y_1 & z_1 & t_1 & 1 \\ x_2 & y_2 & z_2 & t_2 & 1 \\ x_3 & y_3 & z_3 & t_3 & 1 \\ x_4 & y_4 & z_4 & t_4 & 1 \end{vmatrix} = 0$$

The four fixed points must not lie in a plane.

Exercise 4.1

1. $y = -3x - 2$; $y = -\dfrac{x}{3} - \dfrac{2}{3}$

2. $\alpha = \beta + \dfrac{n\pi}{2}$ where $n = 0, 1, 2 \ldots$

3. $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

4. $55k - 55$; $\boldsymbol{x} = 0$; $\boldsymbol{x} = \begin{pmatrix} -2 \\ 1 \\ 3 \end{pmatrix}$

6. $|A| = 3 \ A^{-1} = \begin{pmatrix} -\frac{1}{3} & \frac{11}{3} & -2 \\ 1 & -2 & 1 \\ \frac{2}{3} & -\frac{1}{3} & 0 \end{pmatrix}$

12. $\frac{2}{5}$

Exercise 4.2

2. (i) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$ 6. $\operatorname{Tr} C = 1 + 2\cos\phi$

Exercise 5.1

The answers are given using the theorem on p. 125.

1. $\lambda_1 = 7, \lambda_2 = 2$; $k_n = \frac{1}{5}(7^n - 2^n)$
2. $\lambda_1 = 1, \lambda_2 = -\frac{1}{5}$; $k_n = \dfrac{5}{6}\left(1 - \left(-\dfrac{1}{5}\right)^n\right)$. As $n \to \infty$ $\begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}$
3. $\lambda_1 = \frac{1}{2}, \lambda_2 = \frac{1}{3}$; $k_n = 3((\frac{1}{2})^n - (\frac{1}{3})^n)$
4. $\lambda_1 = 5, \lambda_2 = 2$; $k_n = \frac{1}{3}(5^n - 2^n)$
5. $\lambda_1 = \lambda_2 = 5$; $k_n = n5^{n-1}$
6. $\lambda_1 = \lambda_2 = 4$; $k_n = n4^{n-1}$
7. $\lambda_1 = \lambda_2 = 2$; $k_n = n2^{n-1}$
8. $\lambda_1 = \lambda_2 = -1$; $k_n = n(-1)^{n-1}$
9. $\begin{pmatrix} -3 & -9 \\ 1 & 3 \end{pmatrix}$

Exercise 5.3

1. $\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$ and is independent of p.

2. (i) 0·536 (ii) Probability of a breakdown → 1.
(iii) Probability of correct storage → $\frac{1}{2}$.
(*a*) 0·59049 (*b*) 0·37928

3. 0·375; 0·1765

4. 0·24; 0·072

5. (i) $\frac{1}{2}$ (ii) $\frac{1}{2}$

6. $a_1 = a_2 = a_3 = \frac{1}{3}$

$$\begin{array}{cc} & k \\ & \begin{array}{ccc} S & R & C \end{array} \\ k+1 \begin{array}{c} S \\ R \\ C \end{array} & \begin{pmatrix} \frac{1}{2} & \frac{3}{10} & \frac{2}{7} \\ \frac{1}{3} & \frac{2}{5} & \frac{2}{7} \\ \frac{1}{6} & \frac{3}{10} & \frac{3}{7} \end{pmatrix} \end{array}$$

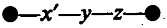
Exercise 6.1
There are a variety of possibilities for each answer.

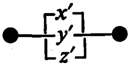
1. $wy' + w'yz + xyz$
2. $z' + x'y' + wx$
3. $xy + x'wz + x'y'wz'$
4. $xyz' + x'y'w' + y'zw' + xz'w + x'yzw$
5. $y'w' + xw'z' + x'y'w$
6. $x'zw + xyzw + x'y'z'w' + xy'zw' + xy'z'w$
7. $w + y'z + x'y'z'$
8. $xy + w'y + w'z$

Exercise 6.2

1. (*a*) $\begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix}$ (*b*) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix}$

(*c*) $\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$ (*d*) $\begin{pmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$





2. (*a*) $\begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}$ (*b*) $\begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$

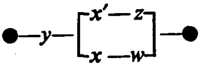
(*c*) $\begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ (*d*) $\begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$

(*e*) $\begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}$ (*f*) $\begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}$

(*g*) $\begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

3. (*a*) $\begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$ (*b*) $\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$ (*c*) $\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$

(*d*) $\begin{pmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$



4. (*a*) $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$ (*b*) $\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix}$ c) $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$

(*d*) $\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix}$

Exercise 7.1

1. (*a*) $\begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ (*b*) $\begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$

2. $\begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$ 3. Yes

4. (*a*) $\begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}$ (*b*) $\begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$

5. $\begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \end{pmatrix}$

6. (*a*) $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ (*b*) $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

(*c*) $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ (*d*) $\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

(*e*) $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ (*f*) $\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$

7. (*a*) $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ (*b*) $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$

8. (*a*) $\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ (*b*) $\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$

Exercise 7.2

1. Cycle $\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and repeating.

2. Cycle $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ and repeating.

3. Cycle $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ repeating to $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

4. Cycle $\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and repeating.

Exercise 8.1

3. $\begin{pmatrix} 0 & 3 & 7 & 8 & 9 \\ 3 & 0 & 4 & 5 & 6 \\ 7 & 4 & 0 & 2 & 3 \\ 8 & 5 & 2 & 0 & 2 \\ 9 & 6 & 3 & 2 & 0 \end{pmatrix}$ 4. *acdefg* 17

7. $A = \begin{array}{c} \\ a \\ b \\ c \\ d \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{pmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \end{pmatrix} \end{array}$; $B = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{ccc} \alpha & \beta & \gamma \end{array} \\ \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \end{array}$;

$$C = \begin{array}{c} \\ a \\ b \\ c \\ d \end{array} \begin{array}{c} \begin{array}{ccc} \alpha & \beta & \gamma \end{array} \\ \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \end{array}$$

9. *abdef* 12 10. *adefgh* 15
11. Truro-Okehampton-Salisbury 166 miles
14. *a*

9.7 $a = -d$

9.13.1 $\begin{pmatrix} u_{n+1} \\ v_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} u_n \\ v_n \end{pmatrix}$

2 $\begin{pmatrix} u_{n+1} \\ v_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} u_n \\ v_n \end{pmatrix}$

Exercise 9.1

1. $u_n = 2^n$ 2. $u_n = 2^n$
3. $u_n = 2^n + 1$ 4. $u_n = 2.4^n + (-3)^n$
6. $u_n = \left(1 - \frac{3\sqrt{7}}{7} i\right)\left(\frac{1}{2} + \frac{\sqrt{7}}{2} i\right)^n + \left(1 + \frac{3\sqrt{7}}{7} i\right)\left(\frac{1}{2} - \frac{\sqrt{7}}{2} i\right)^n$

Exercise 9.2

1. $x = (B - A)\,e^{2t} + (2A - B)\,e^t$
2. $x = (B - 3A)t\,e^{3t} + A\,e^{3t}$
3. $x = \frac{B}{2} \sin 2t + A \cos 2t$

Exercise 9.3

1. $x = \frac{1}{3}t^2 + \frac{5}{9}t + \frac{19}{54}$ 2. $x = \frac{1}{12}(\cos t - \sin t)$

Exercise 9.4

1. $x = e^{2t}(At + B)$ 2. $x = e^{-3t}(Ct + D)$

Exercise 9.5

1. $x = e^{2t}(t^2 + At + B)$ 2. $x = e^t(\frac{1}{2}t^2 + Ct + D)$
3. $x = E\,e^{3t} + e^t\,(-\frac{1}{2}t + F)$ 4. $x = t(G \sin 2t + H \cos 2t)$

Exercise 9.6

1. $\lambda_1 = j\sqrt{2 + \sqrt{2}}$ $\lambda_2 = -j\sqrt{2 + \sqrt{2}}$
$\lambda_3 = j\sqrt{2 - \sqrt{2}}$ $\lambda_4 = -\sqrt{2 - \sqrt{2}}$
Solution as in worked example.

3. $\begin{pmatrix} \cos \omega\theta & -\sin \omega\theta \\ -\sin \omega\theta & \cos \omega\theta \end{pmatrix}$

Exercise 9.10

1. Strictly determined. Value 3. 2. Not strictly determined.
3. Not strictly determined. 4. Not strictly determined.
5. Strictly determined. Value 0. 6. Not strictly determined.

Exercise 9.11

1. $p_1 = \frac{7}{9}$; $q_1 = \frac{4}{9}$; $v = \frac{44}{9}$ 2. $p_1 = \frac{3}{4}$; $q_1 = \frac{1}{2}$; $v = 4\frac{1}{2}$
3. $p_1 = \frac{7}{9}$; $q_1 = \frac{4}{9}$; $v = 7\frac{8}{9}$ 4. Strictly determined. $v = 3$
5. $p_1 = \frac{6}{7}$; $q_1 = \frac{4}{7}$; $v = \frac{4}{7}$ 6. Strictly determined. $v = 4$
7. $p_1 = \frac{2}{3}$; $q_1 = \frac{1}{3}$; $v = \frac{2}{3}$
8. $p_1 = \frac{12}{13}$; $v = \frac{24}{13}$. He should buy the chocolates regardless!

Exercise 9.12

1. 3 2. 3 3. 3 4. 7
5. 9 6. 7 7. 15 8. 21

This book is intended as an extension and continuation of *Matrices 1* and *Matrices 2* (both by G. Matthews in the *Contemporary School Mathematics* series), at sixth form level. The book contains a short revision of 'O' level matrix work but is, in fact, completely self-contained. The first section is an investigation of some aspects of matrix algebra and covers all the known 'A' level syllabuses. The purpose of this section is to provide the mathematical tools, which will be of important use later on when dealing with the applications. The second section gives an insight into some of the many and varied applications of matrices.

It is impossible to include a complete list of applications as this would be too large, but the selection given has been taken from as wide a field as possible and is of the greatest interest.

In preparation

Analogue Computers—
Michael Brand and Timothy Brand
An Introduction to Operational Research—
J. O. Thompson

Edward Arnold

ARNO
£1·30 net

SBN: 7131 2246 3

Brand &
Sherlock

MATRICES: Pure and Applied

Arnold